the CATO JOURNAL

AN INTERDISCIPLINARY JOURNAL OF PUBLIC POLICY ANALYSIS

FEDERALISM AND THE ECONOMIC ORDER

Peter H. Aranson	Federalism: The Reasons of Rules
Aaron Wildavsky	A Double Security: Federalism as Competition
Thomas R. Dye	The Policy Consequences of Intergovernmental Competition
Bruce L. Benson	Interstate Tax Competition, Incentives to Collude, and Federal Influences
Richard Vedder	Tiebout, Taxes, and Economic Growth
Randall G. Holcombe	Growth Management in Florida: Lessons for the National Economy
Bernard H. Siegan	Land Use Regulations Should Preserve Only Vital and Pressing Governmental Interests

OTHER ARTICLES

James D. Gwartney	A Positive Proposal to Improve Our Schools
Charles W. Baird	Labor Law Reform: Lessons from History
Martin Ricketts	Mobility Barriers and Public Policy
Tyler Cowen and Randall Kroszner	Mutual Fund Banking: A Market Approach
Melanie S. Tammen	The Precarious Nature of Sovereign Lending: Implications for the Brady Plan
George A. Selgin	Monetary Equilibrium and the Productivity Norm of Price-Level Policy

VOLUME 10 NUMBER 1 • PUBLISHED BY THE CATO INSTITUTE • SPRING/SUMMER 1990

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the CATO JOURNAL

VOLUME 10 NUMBER 1 • PUBLISHED BY THE CATO INSTITUTE • SPRING/SUMMER 1990

Editor
JAMES A. DORN

Book Review Editor
CATHERINE ENGLAND

Production
DAVID LAMPO

EDITORIAL BOARD

TERRY L. ANDERSON
Montana State University

PETER H. ARANSON
Emory University

CHARLES W. BAIRD
California State University
at Hayward

T. E. BORCHERDING
Claremont Graduate School

YALE BROZEN
University of Chicago

JAMES M. BUCHANAN
Center for Study of
Public Choice
George Mason University

RICHARD A. EPSTEIN
University of Chicago

MARILYN R. FLOWERS
University of Oklahoma

STEVE H. HANKE
Johns Hopkins University

FRIEDRICH A. HAYEK
University of Freiburg

ROBERT HIGGS
Seattle University

JERRY L. JORDAN
First Interstate Bancorp

JOSEPH P. KALT
Harvard University

LEONARD P. LIGGIO
Institute for
Humane Studies

DONALD N. McCLOSKEY
University of Iowa

WILLIAM A. NISKANEN
Cato Institute

SAM PELTZMAN
University of Chicago

ALVIN RABUSHKA
Hoover Institution
Stanford University

VERNON L. SMITH
University of Arizona

ROLAND VAUBEL
Mannheim University

KAREN I. VAUGHN
George Mason University

RICHARD E. WAGNER
George Mason University

LAWRENCE H. WHITE
University of Georgia

LELAND B. YEAGER
Auburn University

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The CATO JOURNAL (ISSN 0273-3072) is published in the spring/summer, fall, and winter by the Cato Institute, 224 Second Street SE, Washington, DC 20003. Annual subscriptions are $21.00 for individuals and $35.00 for institutions. Single copies are $7.00. Two-year and three-year subscriptions are $42.00 and $63.00, respectively, for individuals, and $70.00 and $105.00, respectively, for institutions. Foreign subscribers should add $5.00 per year for regular delivery and $10.00 per year for airmail delivery.

Correspondence regarding subscriptions, changes of address, procuring back issues, advertising and marketing matters, etc., should be addressed to the Publications Department. All other correspondence, including requests to quote or reproduce material, should be addressed to the editor.

Unsolicited manuscripts cannot be returned and will be acknowledged only if accompanied by a stamped, self-addressed envelope. There is no submission fee. Articles should focus on specific public policy issues. Accepted manuscripts must conform to the CATO JOURNAL's style requirements. (A style sheet is available on request.)

Authors should submit three copies of their papers or communications to the editor. An abstract of no more than 250 words should also be included to aid referees. Manuscripts should have wide margins and be double-spaced throughout, including footnotes.

Unsolicited book reviews cannot be returned and will be acknowledged only if they are accepted for publication. They should conform to the above requirements and should provide an in-depth analysis of major studies of public policy. Reviews should range from 800 to 1,000 words.

The views expressed by the authors of the articles are their own and are not attributable to the editor, the editorial board, or the Cato Institute.

The CATO JOURNAL is indexed and abstracted in the *Journal of Economic Literature, Current Contents/Social and Behavioral Sciences, Social Sciences Citation Index, America: History and Life, Historical Abstracts,* and the **ABI/INFORM** database. The CATO JOURNAL is available on microfilm from University Microfilms International, Ann Arbor, Michigan 48106.

Printed in the United States of America.

Copyright © 1990 by the Cato Institute.
All rights reserved.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

CATO JOURNAL • VOLUME 10 NUMBER 1 • SPRING/SUMMER 1990

FEDERALISM AND THE ECONOMIC ORDER

Introduction 1

Peter H. Aranson
Federalism: The Reasons of Rules 17

Aaron Wildavsky
A Double Security: Federalism as Competition 39

Thomas R. Dye
The Policy Consequences of Intergovernmental Competition 59

Bruce L. Benson
Interstate Tax Competition, Incentives to Collude, and Federal Influences 75

Richard Vedder
Tiebout, Taxes, and Economic Growth 91

Randall G. Holcombe
Growth Management in Florida: Lessons for the National Economy 109

Bernard H. Siegan
Land Use Regulations Should Preserve Only Vital and Pressing Governmental Interests 127

OTHER ARTICLES

James D. Gwartney
A Positive Proposal to Improve Our Schools 159

Charles W. Baird
Labor Law Reform: Lessons from History 175

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Martin Ricketts
Mobility Barriers and Public Policy 211

Tyler Cowen and Randall Kroszner
Mutual Fund Banking: A Market Approach 223

Melanie S. Tammen
The Precarious Nature of Sovereign Lending: Implications for the Brady Plan 239

George A. Selgin
Monetary Equilibrium and the Productivity Norm of Price-Level Policy 265

COMMUNICATIONS

Richard B. Coffman
Why Is Education Publicly Provided? Comment on Lott 289

John R. Lott, Jr.
Why Is Education Publicly Provided? Some Further Thoughts 293

BOOK REVIEWS

The Political Economy of the GATT
Peter Moser

Reviewed by **William A. Niskanen** 299

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

INTRODUCTION

FEDERALISM AND THE ECONOMIC ORDER

James A. Dorn

> For the efficient management of your common interests in a country so extensive as ours, a government of as much vigor as is consistent with the perfect security of liberty is indispensable. Liberty itself will find in such a government, with powers properly distributed and adjusted, its surest guardian.
>
> —George Washington[1]

Federalism and the Constitution of Liberty

Constitutional rules help shape the economic order. A constitution is a fundamental law and differs significantly from legislation.[2] The rules embodied in the U.S. Constitution reflect the *principles* that underlie the document. The most basic principle underlying the U.S. Constitution is that the primary function of government is to preserve liberty—that is "to maintain all in the secure and tranquil enjoyment of the rights of person and property," as Washington put it in his Farewell Address (1796). American federalism was designed with this principle in mind.

The U.S. Constitution is often referred to as a "charter of freedom"[3]—or, to use F. A. Hayek's expression, a "constitution of liberty" (Hayek 1960) protecting both personal and economic liberties. To secure "a government of as much vigor as is consistent with the perfect security of liberty," the Framers proposed, and the states consented to, a federalist system of government (a system that Washington thought would preserve liberty, provided governmental powers were "properly distributed and adjusted").

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

[1]Farewell Address, 17 September 1796.

[2]See, for example, Dicey ([1915] 1982) and Hayek (1960, 1982).

[3]See Dorn (1988, especially section V and the references therein).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The choice of a constitutional democracy and a federalist structure of government was no accident. James Madison, the chief architect of the Constitution and the Bill of Rights, devoted many hours to the study of alternative forms of government. He used his broad knowledge to help convince his fellow delegates and countrymen to adopt a system of government in which governmental powers would be limited, the national government would be unable to interfere with the powers of state governments over a wide range of issues, and individual rights would be safeguarded.

The problem of framing a government was best stated by Madison in *Federalist* No. 51:

> If men were angels, no government would be necessary. If angels were to govern men, neither external nor internal controls on government would be necessary. In framing a government which is to be administered by men over men, the great difficulty lies in this: you must first enable government to control the governed; and in the next place oblige it to control itself.

To resolve the difficulty of framing a system of government that would provide an effective mechanism for governing the people and at the same time govern itself, Madison and the Framers turned to a federalist system or a compound republic. It was thought that a system characterized by decentralization and competition would be conducive to the efficient conduct of governmental affairs. It was also thought that consititutional limitations on governmental powers would effectively check the Leviathan's appetite for ever-greater power.

In contrast to a unitary system of government, therefore, a compound republic was thought to offer added security to individual rights. As Madison wrote (*Federalist* No. 51):

> In the compound republic of America, the power surrendered by the people is first divided between two distinct governments, and then the portion allotted to each subdivided among distinct and separate departments. Hence a *double security* arises to the rights of the people. The different governments will control each other, at the same time that each will be controlled by itself [emphasis added].

By making it clear that "powers not delegated to the United States by the Constitution, nor prohibited by it to the States, are reserved to the States respectively, or to the people" (Tenth Amendment), the Constitution gave life to the federalist principle and to the competitive spirit of government favored by the Framers.

Moreover, the Constitution—by limiting the spending and taxing powers of government (Article I, Sections 8–10; Tenth Amendment),

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

by maintaining internal free trade (Article 1, Section 8, the commerce clause; Article I, Section 9), by protecting freedom of contract (Article I, Section 10, the contracts clause), by safeguarding property rights (Fifth Amendment, and later the Fourteenth Amendment), and by recognizing the priority of individual rights (First and Ninth Amendments)—enshrined the principle of freedom in the federalist system and provided the basis for a liberal economic order.

Whether or not the Framers intended to create a private market economy, they in fact did so by safeguarding and nurturing the institutions necessary for its existence—namely, a rule of law, private property, and freedom of contract.[4] As Niskanen (1988, p. xii) has argued:

> Any reading of the deliberations that led to the Constitution and the several amendments should lead one to recognize that these rules were specifically designed as part of the larger set of rules to secure individual rights. The Framers may not have shared a common vision about the economic system. There should be no doubt, however, that the Constitution was designed to provide a strong but limited federal role, free trade among the states, and the security of private property—however the intent of the Framers may have been changed by subsequent interpretation.

The connection between federalism and the economic order is the theme of the first seven papers in this issue.[5] In particular, these papers explore the nature of American federalism, how it has changed over time, and the impact of intergovernmental competition on efficiency and economic growth. Before discussing each of these papers, it is useful to examine more carefully the key features of American federalism with an eye toward the role of the judiciary in a fully developed federalist system.

American Federalism and the Judiciary

In his classic study *The Law of the Constitution*, A. V. Dicey [1915] 1982, p. 77) listed the three key features of a fully developed federalist system, such as existed in the United States: (1) "the

[4]See Lee (1987) who argues that it is largely immaterial whether the framers intended to promote a capitalist economic order. What is important is that they sought to limit the coercive power of government, and in so doing set the basis for a viable market order.

[5]These papers were first presented at the Sixth Annual Critical Issues Symposium—"Competition among Governments, Efficiency, and Economic Growth"—sponsored by the **Florida State University Policy Sciences Program** and the **James Madison Institute**, 8–10 March 1990. Special thanks go to James D. Gwartney for organizing the symposium and for suggesting that selected papers be used in the *Cato Journal*.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

supremacy of the constitution," (2) "the distribution among bodies with limited and co-ordinate authority of the different powers of government," and (3) "the authority of the Courts to act as interpreters of the constitution."

The U.S. Constitution is "the supreme Law of the Land" (Article VI); it distributes and limits the powers of government so that the rights of persons and property are secure. As Dicey (p. 84) notes:

> The United States Constitution prohibits both to Congress and to the separate states the passing of a bill of attainder or an *ex post facto* law, the granting of any title of nobility, or in effect the laying of any tax on articles exported from any State, enjoins that full faith shall be given to the public acts and judicial proceedings of every other State, hinders any State from passing any law impairing the obligation of contracts, and prevents every State from entering into any treaty, alliance, or confederation; thus it provides that the elementary principles of justice, freedom of trade, and the rights of individual property shall be absolutely respected throughout the length and breadth of the Union

Because the U.S. Constitution is the supreme Law of the Land, "Judges in every State shall be bound thereby" (Article VI); and, in a federal system, the Supreme Court necessarily becomes "the final interpreter of the Constitution" (Dicey, p. 89). Thus, Dicey (p. 91) writes: "To the judiciary in short are due the maintenance of justice, the existence of internal free trade, and the general respect for the rights of property."

The Supreme Court's role as "final arbiter" places it in a precarious position; for if the Court succumbs to political pressure, it may end up losing its moral authority as guardian of the Framers' Constitution of liberty. And if the Consitution as a charter of freedom is eroded, either with regard to personal freedom or economic liberty, American federalism will also fail as a limit on the coercive power of government.

It was with this danger in mind that Dicey (p. 101) warned against the politicization of the judiciary and pointed to *Munn v. Illinois* as an indication that

> the most honest judges are after all only honest men, and when set to determine matters of policy and statemanship will necessarily be swayed by political feeling and by reasons of state. But the moment that this bias becomes obvious a Court loses its moral authority, and decisions which might be justified on grounds of policy excite natural indignation and suspicion when they are seen not to be fully justified on grounds of law.

The *Munn* decision (94 U.S. 113 [1877]) is of special importance because with it the Court established the so-called public interest

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

doctrine, which extended the police powers of the states over economic life. As a result, private property rights and freedom of contract were less secure and the legislative door was opened to the rise of the redistributive state. In *The Birth of a Transfer Society,* Anderson and Hill (1980, p. 63) explain the significance of *Munn*:

> Under the doctrine of police power, the Supreme Court had previously upheld the power of the states to regulate economic activity incidentally. But the idea that private property used in the public interest was subject to public regulation gave a broad new authority to the government to interfere with contracts and private property. Such authority raised the returns to transfer activity by giving nonowners the potential to affect resource allocation.

In his dissenting opinion, Justice Field recognized the threat *Munn* and the public interest doctrine posed for the survival of American federalism and private property rights. Echoing the Framers' constitutional principles and and their original understanding of federalism and its link to economic liberty, Field (at 140) stated:

> If this be sound law, if there be no protection, either in the principles upon which our republican government is founded, or in the prohibitions of the Constitution against such invasion of private rights, all property and all business in the State are held at the mercy of a majority of its legislature.

State regulation over economic life continued with the decision in *Mugler v. Kansas* (123 U.S. 623 [1887]), where the Court upheld a state ban against the sale of liquor. However, in rendering his opinion, Justice Harlan reflected on the limitations of the police powers and paved the way for extending substantive due process protection to economic liberties via the Fourteenth Amendment.[6] According to Justice Harlan (at 661):

> There are, of necessity, limits beyond which legislation cannot rightfully go. [T]he courts must . . . upon their own responsibility, determine whether, in any particular case, these limits have been passed. . . . If, therefore, a statute purporting to have been enacted to protect the public health, the public morals, or the public safety has no real or substantial relation to those objects, *or is a palpable invasion of rights secured by the fundamental law,* it is the duty of the courts to so adjudge, and thereby to give effect to the Constitution [emphasis added].

When the Court gives effect to the Framers' Constitution of liberty, it also gives effect to American federalism—both as a procedure for competitive government and as a doctrine of substantive limits on

[6]See, for example, Lockhart, Kamisar, and Choper (1970, p. 455).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

governmental powers. The federalist system was designed not merely to decentralize governmental functions but to restrict both federal and state powers so that the sanctity of individual rights would be maintained.[7]

The *structure* of U.S. federalism is the distribution of powers, but the *substance* of federalism is freedom. If the Court fails to protect the substance of federalism, mere decentralization will not be sufficient to protect the rights of individuals. And if economic freedom is not afforded the same protection as personal freedom, the market order will weaken as federalism is undermined by an expansion of the police powers both at the national level and at the state level.

Applying substantive due process to both civil and economic liberties is entirely consistent with the Framers' view of federalism and with the principles underlying the Constitution as fundamental law.[8] Whether they like it or not, judges must interpret the Constitution. Their decisions will spell the difference between a federalism consistent with a liberal economic order and a federalism that upsets the spontaneous market process by interfering with the property right—broadly defined in the Madisonian tradition as "every thing to which a man may attach a value and have a right: and *which leaves to everyone else the like advantage*" (Madison [1792] 1906, p. 101).

Under the rubric of property, Madison would include an individual's right to "the safety and liberty of his person" as well as an individual's rights to "the free use of his faculties and free choice of the objects on which to employ them" (Madison, p. 101). Since the function of the judiciary, in Madison's view, is to protect individual rights, it should act as "an impenetrable bulwark against every assumption of power in the Legislative or Executive" branch (*Annals* 1: 439). "The Federal judiciary," wrote Madison ([1833] 1865, pp. 296–97), "is the only defensive armour of the Federal government, or, rather, for the Constitution and laws of the United States. Strip it of that armour, and the door is wide open for nullification, anarchy, and convulsion." In short, the judiciary is the main pillar upon which American federalism stands or falls.

[7]In this sense, fiscal federalism is only a subset of federalism as originally understood by the Framers. Moreover, as Bish (1987, pp. 386–87) points out, fiscal federalism, by focusing on problems of scale and ignoring the monopoly problem, runs the risk of losing the competitive character that distinguishes it from a decentralized *unitary* system of government.

[8]See, for example, Siegan (1980) and the essays in Dorn and Manne (1987); also see Dorn (1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Erosion of Economic Liberties and Transformation of Federalism

With the termination of substantive due process protection for economic liberties in the late 1930s and the Court's broad interpretation of the commerce clause (to encompass federal regulation of almost any economic activity whether or not it was specifically related to interstate commerce), the original conception of American federalism was turned on its head.[9] As *Peter H. Aranson* shows in his paper, federalism was transformed through a series of Court decisions from meaning a legal doctrine of "constitutional decentralization" to meaning "contingent decentralization"—with the presumption that Congress would decide the distribution of powers, regardless of the impact on private rights.

Thus, instead of being driven by the Framers' vision of a national government limited to those powers enumerated in the Constitution, American federalism has become dominated by the modern vision of a national government that defines its own powers. In the process, the redistributive state has risen to occupy the territory that was once governed by a more limited state. Moreover, as *Aaron Wildavsky* points out, the modern vision of egalitarianism has tended to crowd out the Framers' vision of a self-regulating system of competitive federalism. In his view, egalitarianism and competition are ultimately irreconcilable.

Constitutional Decentralization versus Contingent Decentralization

Aranson traces the process by which American federalism, as a doctrine limiting the overall powers of government and protecting individual rights, has been undermined by judicial decisions that have lost touch with the Framers' principled approach to federalism and its grounding in a Constitution of liberty. Specifically, he shows how, paradoxically, the Court's broad interpretation of the commerce clause has allowed the federal government to regulate *intra*state trade while at the same time allowing states to interfere with *inter*state commerce. In effect, the Court has widened both federal and state regulatory powers and has eroded the economic Constitution of liberty without any formal amendment process.[10]

Unlike "constitutional decentralization," in which decentralization of powers is strictly limited by the Constitution as a charter

[9]The demise of substantive economic due process is discussed in Siegan (1980). The Court's treatment of the commerce clause is discussed in Epstein (1987).

[10]On the erosion of the economic Constitution, see Niskanen (1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of freedom, "contingent decentralization" allows purely utilitarian calculus and majoritarianism to determine the distribution of powers as well as the scope of government action. As such, the true nature of federalism is destroyed and with it the benefits of intergovernmental competition that flow from limited government and a proper distribution of powers.

Aranson's message is that contingent decentralization alone is insufficient to prevent political forces from leading to the recentralization of governmental power. What is necessary is that the judiciary fulfill its role as guardian of the Framers' Constitution and ultimately that the people retain the liberal republican virtues that were evident at the founding of our compound republic.[11]

Competitive Federalism versus Egalitarianism

Until the mid-1930s, the United States remained largely a protective rather than a redistributive state. The limitation of the government's redistributive function was in line with a strict interpretation of the enumerated powers of the federal government, which, as Niskanen (1985, p. 2) notes, provide "*no* explicit authority for federal welfare programs."

The termination of substantive economic due process after 1936, the Court's ruling in *United States v. Butler* (297 U.S. 1 [1936])[12] that effectively removed constitutional constraints on the scope of federal spending by adopting a so-called public-purpose test, the change in popular sentiment toward redistribution as a by-product of the Great Depression, and the rise of modern "democratic liberalism" all helped spur the rise of the redistributive state or what Anderson and Hill (1980) have called the "transfer society."[13]

The transition to the modern welfare state has been accompanied by a transition from a rule-of-law approach to equality—that is, a concern with ensuring equal protection under just laws without

[11]For Madison, liberal republicanism essentially meant the establishment of "the efficacy of popular charters, in defending liberty against power . . . and in keeping every portion of power within its proper limits" (as discussed in his essay "Charters" [1792]).

The spirit of Madison's liberal republicanism is captured in the following passage from John O'Sullivan, who in 1837 wrote:

> The best government is that which governs least. . . . [Legislation] should be confined to the administration of justice, for the protection of the natural equal rights of the citizen, and the preservation of the social order. In all other respects, the voluntary principle, the principle of freedom . . . affords the true golden rule [in Vernier 1987, pp. 12–13].

[12]For a discussion of this case and its implications for the welfare state, see Niskanen (1985, p. 6).

[13]For a general discussion of the transfer society and the conditions surrounding its rise since the 1930s, see Dorn (1986).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

regard to outcomes—to a purely political approach to equality in which legislation is used to alter the "rules of the game" to achieve more equal outcomes. Whether the agenda of rent-seeking groups interferes with private property rights is of secondary importance to the partisans of the welfare state, whose declared aim is "social justice" (defined as some desired income distribution, not as a state of affairs consistent with the rule of law).[14]

What are the implications of this change in the concept of equality for American federalism? According to Wildavsky, the shift toward the redistributive state spells the demise of federalism as a system of intergovernmental competition designed to encourge experimentation and to test alternative arrangements for the provision of public goods and services.

Competition means diversity. As such, it is inconsistent with egalitarianism, argues Wildavsky. One of the main benefits of intergovernmental competition is the increased efficiency resulting from the closer link between taxpayer preferences and governmental response. The benefits from increased efficiency, however, will be dissipated if "equity" considerations (in the sense of redistributive justice) become predominant.

Thus Wildavsky warns that the emergence of the modern welfare state, with its focus on equal outcomes, has tended to preclude competitive federalism. Moreover, with the increasing centralization of governmental functions at the federal level, the "double security" that Madison spoke of has also been weakened, with adverse consequences for private rights. What concerns Wildavsky is not simply the change in federal-state relations but rather the weakening of "the self-regulatory principles" that constitute the essence of Madisonian federalism.

Intergovernmental Competition, Collusion, and Economic Growth

The papers by *Thomas R. Dye*, *Bruce L. Benson*, and *Richard Vedder* discuss various facets of competitive federalism: the ability of intergovernmental competition to satisfy taxpayer-consumer preferences for efficiency and equity, the incentive for governments to collude in order to increase their revenues, and the importance of differences in state tax rates for economic growth.

[14]On the inconsistency of redistributive or social justice with a free market order and a rule of law, see Hayek (1982, chaps. 8–9). On the changing concept of equality, see Dorn (1990).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Policy Outcomes under Intergovernmental Competition

Under a fully developed system of federalism, state and local governments must compete for their customers—the taxpayer-consumers who have the option of "voting with their feet" (see Tiebout 1956). Voters, however, are interested in both efficiency and equity—with equity being viewed in terms of end-states rather than in terms of just procedures or processes.[15] Thus, as Dye points out, taxpayer-consumers may wish to trade off some additional wealth for greater equality of income.

In the real world of politics, policy preferences will help shape constitutional interpretation. For this reason Dye argues that "debates over federalism are only lightly camouflaged debates over policy." While Wildavsky is concerned with the adverse effect of egalitarianism on the federalist principle of self-regulation and on the social benefits flowing from competition among governments, Dye notes that if taxpayer-consumers have a preference for the redistributive function of government (in addition to allocative efficiency), then policymakers will respond to that preference in a politically motivated federalist system.

In short, "Competitive federalism ensures responsiveness to redistributional preferences, as well as other policy preferences," writes Dye. The problem, of course, is that adherence to federalism as contingent decentralization jeopardizes the Framers' original understanding of federalism as constitutional decentralization, to use Aranson's terminolgy,

The Incentive to Collude

Individuals in both the private sector and in the public sector have an incentive to collude. Collusion, however, is unlikely to be successful unless the parties to the agreement can be effectively monitored and the collusive agreement enforced. In the private sector, such collusion is often accomplished by government grants of privilege in the form of licenses, quotas, and other restrictive policy measures. In the public sector, collusion among state governments to mute the effects of tax competition is often fostered by federal assistance in the form of deductibility of state and local taxes and by revenue sharing. But as Benson shows, the benefits gained by politicians from lessening tax rate competition are dissipated, in part, by resorting to non-tax competition. "Incentives to compete," writes

[15]For a discussion of the difference between end-state and procedural justice or equity, see Nozick (1974, especially pp. 149–82).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Benson, "do not disappear in the face of a lessening of the effectiveness of one particular avenue of competition."

Federalism, Tax Competition, and Growth

The tax regime is an important part of any economic order or regime. Under a federalist structure of government, taxpayer-consumers have greater freedom to respond to tax increases than under a unitary system of government. This greater freedom of choice stems from the fact that under federalism taxpayer-consumers have two options: the "exit" option and the "voice" option.[16] If state A increases tax rates relative to state B, then at the margin some individuals may decide to leave state A and enter state B, while others may voice their objections to the higher tax rates and use the democratic process to restore the lower rates. Together, both voice and exit options provide taxpayer-consumers with greater protection against the tyranny of government than exists in a unitary system in which only the voice option exists (unless one wishes to leave the country, a very high-cost option).

Given the options to exit a high-tax state and to voice opposition through the voting machine in a federalist system, two questions come to mind: (1) Are people responsive to tax-rate changes? (2) Do states with higher tax rates have relatively lower growth rates than low-tax states, assuming other things remain constant?

Vedder addresses these questions and finds that as taxes increase, people become more sensitive to further tax increases. He also finds that tax-rate differentials do matter in determining differences in interstate growth rates.

The Regulatory State

If liberty is to be preserved in a federalist system, property rights must be respected and enforced. A primary function of government is the safeguarding of private property rights, so that the market process rather than the political process can regulate economic life. As government moves away from protecting property rights to redistributing rights via direct takings as well as through the regulatory process, respect for private property will lessen. The resulting uncertainty, in turn, will hamper the rational use of resources.

The politicization of property rights and the misdirection of resources that occurs when the government fails in its primary function—that is, "to maintain all in the secure and tranquil enjoyment

[16]The terms "exit" and "voice" were first used by Hirschman (1970). For a discussion of their significance in a federalist system, see Rasmussen (1987, pp. 397–98).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of the rights of person and property"—is nowhere better exemplified than in the spread of statewide limits on private economic development and state land use regulations.

The papers by *Randall G. Holcombe* and *Bernard H. Siegan* examine the regulatory state in its attempts to control the development process and to restrict the private use of land. They provide further evidence that regulations aimed at the public good often end up violating private rights, impeding economic progress, and having exactly the opposite effects of those intended when the regulations were enacted.

Florida's Growth Management Policy

Holcombe discusses Florida's 1985 Growth Management Act and its lessons for the national economy. The 1985 Act attenuated the land use rights of private owners and tended to shift decisionmaking authority for land use from local governments to the state government. As Holcombe notes, Florida's move toward *statewide* land use planning parallels the trend occurring in other states.

What Holcombe finds is that Florida's attenuation of private land use rights and the centralization of land use decisions have politicized the development process in Florida and resulted in inefficient land use. Because some of the private owner's land use rights now become part of the public domain, these rights are treated as common property. As such, politics rather than economics will drive the development process. One should therefore expect less-efficient use of resources once private property rights become subject to the hand of the state. Moreover, if owners expect their future land use rights to be less secure under Florida's Growth Management Act, they will have an incentive to accelerate development—pushing to complete projects now rather than to wait and have their plans frustrated by state planners.

The lesson of Holcombe's analysis is clear: The rational use of resources requires well-defined *private* property rights. When the state lessens the discipline of market forces on development, one should not expect political forces to produce a superior outcome in terms of either efficiency or justice (in the sense of protection of private rights). As experience with the Growth Management Act indicates, the results of regulation are often very different from those intended.

The Case of Zoning

The general implications of Holcombe's study of Florida's growth management practices for efficiency and justice apply equally to the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

case of zoning, as Siegan shows in his paper. Drawing on a comparison of Houston (a zoning-free city) with Dallas (where zoning laws are enforced), Siegan presents evidence indicating the superiority of the zoning-free approach to urban development. Like Holcombe, he points to how the zoning process politicizes land use and urban development and leads to an inferior use of scarce land compared to a market-driven approach.

In Siegan's view, moving zoning and land use decisions to the federal level would only worsen the inefficiency already experienced at the state and local levels. As the title of his paper indicates, Siegan believes that "land use regulations should preserve only vital and pressing governmental interests." Thus, the presumption in decisions about land use and development should be in favor of open, competitive markets and against master planning and political decisionmaking. "The great lesson of our times," writes Siegan, "is that the forces of production, conservation, and creativity rest principally in the marketplace and not in government."

The Future of Federalism and the Market Order

In his Farewell Address, Washington stated:

> If in the opinion of the people the distribution or modification of the constitutional powers be in any particular wrong, let it be corrected by an amendment in the way which the Constitution designates. But let there be no change by usurpation; for though this in one instance may be the instrument of good, it is the customary weapon by which free governments are destroyed.

Without using the formal amendment process, the path of federalism in the United States has been changed. Instead of adhering to a well-defined path directed toward the preservation of liberty and property as the guiding principle for government action, American federalism has moved to a new path, one mapped out by utilitarian considerations with little regard for the effects of government action on the proper distribution of powers or their limitations and on individual rights.

The move from constitutional decentralization to contingent decentralization has been facilitated, in large part, by a judiciary that has lost sight of the inseparability of liberty and property. As the Supreme Court has deferred to legislative activism in the field of property rights and economic regulation, there has been an increased use of both federal and state police powers to directly and indirectly take private property whenever politically feasible—whether or not it is for a legitimate public use. The contracts clause and the com-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

merce clause have suffered fates similar to the takings clause of the Fifth and Fourteenth Amendments. In particular, the Court has turned the commerce clause on its head so that there are virtually no limits to federal control of state economic activity, and states are engaged in taxing and regulatory acts that impede interstate trade. These changes have altered the Framers' conception of federalism and upset the liberal economic order, which depends on the protection of private property rights and freedom of contract.

The future of federalism in the United States will depend on the restoration of the Framers' Constitution of liberty as the supreme Law of the Land, the reestablishment of the *proper* distribution of powers, and the will of the Court to provide the same protection for property rights as for other civil liberties.

The collapse of communism and central planning has vividly illustrated the close connection between political reform and economic reform or, more specifically, between the political/constitutional order and the economic order. If the liberal market order is to survive, private property and freedom of contract must also survive.

Experience has shown that the federalist system of government, limiting the overall power of government and distributing power so as to enlarge taxpayer-consumer choice and protect individual rights, is an important part of the environment necessary for freedom and prosperity. Most important, U.S. experience with federalism shows that the demise of a truly federalist system is likely to be accompanied by an erosion of the liberal economic order.

References

Anderson, Terry L., and Hill, Peter J. *The Birth of a Transfer Society.* Stanford, Calif.: Hoover Institution Press, 1980.

Annals of the Congress of the United States. Washington, D.C.: Gales and Seaton, 1834.

Bish, Robert L. "Federalism: A Market Economics Perspective." *Cato Journal* 7 (Fall 1987): 377–96.

Dicey, A. V. *Introduction to the Study of the Law of the Constitution.* 8th ed. 1915. Reprint. Indianapolis, Ind.: Liberty Classics, 1982. (The first edition appeared in 1885.)

Dorn, James A. "The Transfer Society." Introduction. *Cato Journal* 6 (Spring/Summer 1986): 1–17.

Dorn, James A. "Public Choice and the Constitution: A Madisonian Perspective." In *Public Choice and Constitutional Economics*, pp. 57–102. Edited by James D. Gwartney and Richard E. Wagner. Greenwich, Conn.: JAI Press, 1988.

Dorn, James A. "Equality: A Constitutional Perspective." In *The Political Legitimacy of Markets and Governments*, pp. 71–103. Edited by Thomas R. Dye. Greenwich, Conn.: JAI Press, 1990.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Dorn, James A., and Manne, Henry G., eds. *Economic Liberties and the Judiciary*. Fairfax, Va.: George Mason University Press, 1987.

Epstein, Richard A. "The Proper Scope of the Commerce Power." *Virginia Law Review* 73 (November 1987): 1387–1455.

The Federalist Papers: A Contemporary Selection. Abridged and edited by Lester DeKoster. Grand Rapids, Mich.: William B. Eerdmans, 1976.

Hayek, Friedrich, A. *The Constitution of Liberty*. Chicago: University of Chicago Press, 1960.

Hayek, Friedrich A. *Law, Legislation, and Liberty*. Reprint (3 vols. in 1). London: Routledge and Kegan Paul, 1982. (Vol. 1, *Rules and Order*, chaps. 1–6; Vol. 2, *The Mirage of Social Justice*, chaps. 7–11; Vol. 3, *The Political Order of a Free People*, chaps. 12–18).

Hirschman, Albert O. *Exit, Voice, and Loyalty*. Cambridge: Harvard University Press, 1970.

Lee, Dwight R. "The *Calculus of Consent* and the Constitution of Capitalism." *Cato Journal* 7 (Fall 1987): 331–36.

Lockhart, William B.; Kamisar, Yale; and Choper, Jesse H. *Constitutional Law*. 3rd ed. St. Paul, Minn.: West Publishing Co., 1970.

Madison, James. "Property." *National Gazette*, 29 March 1792. In *The Writings of James Madison*. Vol. 6: 1790–1802, pp. 101–3. Edited by Gaillard Hunt. New York: G. P. Putnam's Sons, The Knickerbocker Press, 1906.

Madison, James. "Charters." *National Gazette*, 18 January 1792. In *The Papers of James Madison*. Vol. 14: 6 April 1791–16 March 1793, pp. 191–92. Edited by Robert A Rutland et al. Charlottesville: University Press of Virginia, 1983.

Madison, James. "[Letter] to J. C. Cabell." 1 April 1833. In *Letters and Other Writings of James Madison*. Vol. 4: 1829–1836, pp. 296–97. Published by Order of Congress. Philadelphia: J. B. Lippincott and Co., 1865.

Niskanen, William A. "A 'Constitutional' Perspective on Social Welfare Policy." Paper presented at the Ford Foundation Conference on "Social Welfare Policy and the American Future," Racine, Wisc., 3–5 November 1985.

Niskanen, William A. "The Erosion of the Economic Constitution." Foreword. In *Public Choice and Constitutional Economics*, pp. xi–xiii. Edited by James D. Gwartney and Richard E. Wagner. Greenwich, Conn.: JAI Press, 1988.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic Books, 1974.

Rasmussen, David W. "Federalism from a Market Perspective." *Cato Journal* 7 (Fall 1987): 397–402.

Siegan, Bernard H. *Economic Liberties and the Constitution*. Chicago: University of Chicago Press, 1980.

Tiebout, Charles M. "A Pure Theory of Public Expenditures." *Journal of Political Economy* 64 (October 1956): 416–24.

Vernier, Richard. "Interpreting the American Republic: Civic Humanism vs. Liberalism." *Humane Studies Review* 4 (Summer 1987).

Washington, George. Farewell Address, 17 September 1796.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FEDERALISM: THE REASONS OF RULES

Peter H. Aranson

As the next millenium draws near, problems of freedom, governance, and public policy around the globe increasingly have become problems of federalism. In the Soviet Bloc the historically identifiable nations of Eastern Europe have moved from colonial status to an emerging independence. Even in the Soviet Union itself, the Baltic states and separate Soviet republics now seek to loosen the bounds of central control. The nations of Western Europe are stumbling toward greater centralization, by constructing the confederation promised by Europe 1992 (Aranson 1989). Canada struggles with the federal terms of its constitution. And we in the United States, as we have done for nearly all of our national experience, interpret our political and social conflicts in terms of our federal arrangement.

In this universe of discourse, the problem of establishing "models" of federalism, and then of discerning each model's advantages and drawbacks, provides the core of argument. This essay is no different in that regard. So here I first explain that federalism before the American Constitution differed from the Constitution's federalism. And the Constitution's federalism differs from our contemporary understanding of what federalism requires. Early federalism was a matter of individual or personal virtue. The Constitution's federalism, which, following Martin Diamond (1969), I shall call *constitutional decentralization*, was a matter of legal doctrine. And our present understanding of federalism, which I shall call *contingent decentralization*, is wholly utilitarian.

I shall argue, however, that utilitarian federalism cannot survive, for reasons of ignorance and politics. I then explore the Supreme

---

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor and Chairman, Department of Economics, Emory University. This paper summarizes, in part, the arguments in Aranson (1990). But it focuses more clearly on the role of doctrine in constitutional discourse about federalism. The title of the paper derives from Brennan and Buchanan (1985).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Court's shifting interpretation of federalism from doctrine to utilitarianism, to show that indeed, a utilitarian interpretation of federalism ultimately implies its demise. I then argue in closing that only a doctrinal interpretation and defense of federalism can secure its future and, therefore, its personal and utilitarian virtues.

Background to American Federalism

The term "federalism" comes from the Latin root "foedus," meaning "covenant" (Elazar 1987, Ostrom 1989). The "federative" powers of a state, as Locke called them, were those given to rulers, to form agreements with the heads of other states, to ensure nonaggression and mutual protection. So before its American development, federalism referred more to a form of international relations (Diamond 1969), and less to a constitutive organization of local governments or territories under a central power.

This early construction of federalism invoked an ancient concern about the state's appropriate size. For the Greeks a state must be self-contained and sufficiently small to allow each citizen meaningfully to partake in its political life, thereby gaining a full measure of self-actualization. "Man," Aristotle said, "is a political animal," by which he meant not to describe people as they are, but as they could become if the state were small enough to allow for full political participation and, therefore, the achievement of the *telos* or goal of human nature.

By Montesquieu's time, concern for the state's appropriate size had shifted in two ways. First, for a state to survive, it had to be somewhat larger than the Greek *polis*. But second, in the new, larger republic, size must remain limited not for purposes of *internal* self-actualization, but to achieve the *external* benefits of individual "republican virtue," without which the state also might perish. For Montesquieu (1949, p. 126), the essential tradeoff became: "If a republic be small, it is destroyed by a foreign force; if it be large, it is ruined by an internal imperfection."

The Framers of the American Constitution remained impressed with Montesquieu's two-fold argument. But again the core justification had shifted with respect to its ultimate objects. First, the Framers were not content to rely on republican virtue, for they did not believe in its constancy. "The latent causes of faction are . . . sown in the nature of man," said Madison (*Federalist* No. 10). Humanity's history provided ample evidence of "little, jealous, clashing, tumultuous commonwealths," offering the spectacle of "the wretched nurseries of unceasing discord," wrote Hamilton (*Federalist* No. 9). And there-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

fore, as Madison summed up, "ambition must be made to counteract ambition" (*Federalist* No. 51).

Second, the Framers' principal concern remained with preserving the former colonies against foreign aggression (Riker 1964, chap. 2; 1987; n.d.). The peace with Great Britain was an uneasy one, with British forts strategically located in the Northwest Territories. Spain held the Southwest and some of the South, and Kentucky actually considered joining Spain. "The Country," said Gouvernor Morris, "must be united. If persuasion does not unite it, the sword will" (Farrand 1937, p. 530).

It is a close question whether the Framers, and particularly Madison and Hamilton, in constructing their peculiarly American version of federalism, yielded to logical argument or to political necessity. Logical argument recommended Montesquieu's version of federal relations—a confederation—among reasonably small republics, brought together to ward off foreign design. The tenacity with which the states held to their independence and sovereignty likewise recommended such a course from political necessity.

But a strong case had been made against the Articles of Confederation, which had embraced Montesquieu's model of federalism. First, several of the state governments appeared to be internally disordered, given to democratic, legislative excesses resulting in the abrogation of contracts and (to the same effect between debtors and creditors) the inflation of state-issued currencies. Second, interstate trade wars seemed to be real possibilities, as particular states allegedly took actions that weakened the continental common market. Third, the want of a robust national taxing power threatened the states' abilities to raise and maintain armies and navies for mutual protection.

As overdrawn as this case seems in retrospect (Jensen 1940; Kitch 1981; Aranson 1989), it nevertheless suggested a stronger form of union, especially where the federative, military, and treaty-making powers were concerned. But the stubborn independence of the states remained an unchangeable reality with which the Framers had to contend. The result of these forces was a peculiarly American invention, not since duplicated in any large state, an invention that Diamond (1969) called "constitutional decentralization."

Contingent versus Constitutional Federalism

To grasp the full nature of this invention, we must set aside the usual dimension along which we ordinarily measure federalism, namely, that running from decentralization to centralization. Certainly, we can identify the change that the Constitution of 1789

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

wrought in the new nation's position on this dimension: Governance became relatively more centralized. Just as certainly, most Americans, and nearly all economists (Oates 1977) today regard federalism as being merely some degree of decentralization. But that characterization obscures, indeed misconstrues, the invention's real nature. Under constitutional decentralization the degree of centralization (or decentralization) was fixed in particular ways by the law of the Constitution, reserving permanently to the states a measure of sovereignty wholly absent in today's definitions of federalism.

Stated differently, in today's discourse the degree of centralization or decentralization in a federalism locates all sovereignty in the central government. That government then decides, as a matter of prudential or political judgment, how much authority to devolve to the constituent units. The closest version we have to this concept of federalism today occurs in the relations of cities and counties to state governments, which control not only the extent and limits of local government authority, but sometimes even their very boundaries. The result is a degree of decentralization contingent on the central power's choices.

Federalism as *constitutional* decentralization differs from federalism as *contingent* decentralization in that the authority of the states under constitutional decentralization is guaranteed as a matter of organic, constitutional law. Neither prudential nor political judgments or decisions taken at the national level can overturn such guarantees in the face of appropriate legal fidelity to the original constitutional arrangement. To grasp the full implications of this difference between federalism as constitutional decentralization and federalism as contingent decentralization, we must compare the two concepts in operation, under two competing hypotheses.

Federalism and Public Interest

The first hypothesis—the public-interest hypothesis—is that most familiar to economists in writings about the state in general, and about federalism in particular. The core notion holds that governments emerge in human communities to address a set of problems that market relations may find difficult to resolve. These include the definition and protection of rights in person, property, and contract, the production of public goods, the suppression of "public bads," the provision of external benefits, and the regulation of market power.

When applied to the construction of a federal arrangement, the public-interest hypothesis raises three distinct, though overlapping considerations. The first concerns the level at which one or more of these problems occurs. For example, the "catchment" of national

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

defense or international treaty making may be the entire nation, requiring that all or most such decisions be taken at the national level. But the catchment of most road construction, or crime or property law, may be confined to smaller units, recommending that all or most such decisions be taken at the local level. In short, the efficient level and pricing of one or more of these activities may best be chosen at different jurisdictional levels.

Failure to get this matter right can have unfortunate consequences. For example, if taxpayers nationally must pay for Atlanta's public transportation system, then Atlanta will spend too much for its system. And if, by contrast, decisions are made solely in Ohio concerning air pollution generated there but afflicting Eastern states, then Ohio will not adopt the "correct" level of pollution abatement. The general rule is that people must bear the costs of that from which they benefit, and that they must have a right to avoid, or to gain compensation for, the costs that others impose on them.

The second consideration involves economies and diseconomies of scale. In theory, one can expand indefinitely the catchment of a producer of a pure public good, thereby reducing the cost of output per person served. But such goods rarely occur in nature. National defense may provide an imperfect example. A public statue might inspire those who view it. But as the size of the audience increases, crowding and other local aspects become more readily apparent.

Economies and diseconomies of scale more directly affect the production process itself. The optimal size of a school district, for example, is probably much smaller than that which most jurisdictions adopt. Optimum scale occurs at the output where average total cost is minimized, which in turn is the output at which the change in the economies of scale is exactly equal (but of opposite sign) to the change in the diseconomies of scale. Jurisdictional boundaries, therefore, should be adjusted until each jurisdiction provides its output at optimal—least cost—levels.

The third consideration involves what economists call "Tiebout forces" (Tiebout 1956). Getting people to reveal their preferences for public goods or services that exhibit jointness of production and supply is always problematical. People can report that a particular production level is worth little to them, and thereby avoid taxation based on a benefit principle. This is just another variation of the free-rider problem. In addition, to the extent that production becomes centralized, few competitive forces remain to ensure the discovery and adoption of efficient methods of supply.

The Tiebout hypothesis envisions constructing competing jurisdictions and allowing people to reveal their preferences by "voting

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

with their feet." The underlying logic is that people will sort themselves out among jurisdictions that best provide the mix of goods and services that they prefer. Improvements in governmental performance will be capitalized into land prices, thus giving citizens and public servants incentives to act in appropriate ways. Other jurisdictions will emulate successful innovations and avoid failures. In short, the Tiebout hypothesis supposes that it is possible to "make a market" in the competitive supply of local public goods and services.

The public-interest hypothesis thus entails these separate considerations, carried on at different levels of analysis. First, the public sector will confine itself to perform all and only those tasks that have public-interest credentials. Second, the central government will adjust the boundaries and empowerments of lower levels of (subordinate) governments, to accomplish those tasks with a due regard for catchment, scale, and competitive forces. This is what federalism as contingent decentralization means. Stated differently, the national government decentralizes decisionmaking in precisely the same way as would a private sector firm, with the understanding that the full authority to decide on the degree of decentralization remains with the central power.

Federalism as constitutional decentralization tends to disregard these welfare-promoting reasons for decentralization and instead adopts a decentralization as a matter of right and, consequently, of legal doctrine. The difference between constitutional and contingent decentralization, therefore, is that the virtues of decentralization happen, if at all, by design under the first and by felicitous accident under the second.

To the extent that American federalism entails constitutional decentralization, however, we might ask why it is ever a better idea than contingent decentralization. The answer lies partly in the central government's inability to get the tasks and boundaries right. Planning tasks and jurisdictions requires all of the information that would be required to plan an economy itself. That information, as Hayek (1945) and others have argued, is widely and radically decentralized among the citizenry. Hence, it seems doubtful that the central planning-economic calculation problem, which no one has solved with respect to production, could be solved with respect to the far more difficult problems of public goods, services, and jurisdictions.

But we can claim more than that the central government's decisions in these respects simply remain wrong, and probably random. Instead, even under the assumptions of benevolence that characterize the public interest hypothesis, it seems likely that the central

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

government's information will be biased systematically toward appreciating the *benefits* of increased centralization and the *costs* of increased decentralization. Central decisionmakers, after all, will know their own costs and benefits, but not those of decentralized decisonmakers, including the subject citizens themselves. This is simply another application of the theory of "rational ignorance" (Aranson 1989/1990).

Federalism and Private Interest

The second hypothesis holds that governments are concerned less with the allocative chores of the public interest hypothesis and more with the distributive consequences of public policy: Who gets what? This is the rent-seeking model of the political process (Tullock 1967; Krueger 1974; Stigler 1971; Peltzman 1976; Aranson and Ordeshook 1985; Weingast, Shepsle, and Johnsen 1981). Its earliest development entailed the notion that government treats firms in a competitive industry as a cartel, regulating output and entry accordingly. A welfare—deadweight—loss results, along with the production of rents that the political process itself dissipates. Hence, politics becomes wasteful, welfare-degrading, and not at all involved in the processes that the public interest model would contemplate.

Macey (1990) recently has applied this model to the problem of federalism. In his interpretation, government finds it easier to provide rents from regulating some activities at particular jurisdictional levels than at others. For example, if individual state governments sought through compulsory union membership to cartelize labor markets, those states that did not follow this course would attract and keep lower-cost producers to the disadvantage of producers in unionized, high-cost states. Hence, under the rent-seeking model labor-market cartelization should (and does) occur at the national level. Consumers in markets for medical care and legal services, by contrast, are relatively more immobile than are consumers in ordinary product or factor markets. So, local regulation, more easily than federal regulation, can (and does) maximize rents by taking cognizance of local conditions concerning these services.

Changing technology alters the preferred locus of political control in both the public interest and private interest, rent-seeking models. For example, advances in data processing and transmission make it possible for banking and securities transactions increasingly to occur over ever-wider geographical areas. The locus of concerns for financial stability will expand in the public interest model, and the locus of concerns for rent provision will likewise expand in the private interest model.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The normative conclusions derived from each model in the face of such technological changes, however, remain wholly different. If we ignore the problem of information and its bias, we would prefer the superior responsiveness of federalism as contingent decentralization if we believed that government acts principally in accord with the public interest model. But we would prefer the "rigidity" of doctrine associated with federalism as constitutional decentralization if we believed that government acts principally in accord with the private interest model. Knowing which model applies thus becomes crucial for uttering a preference for one form of federalism over the other. If both models find occasional application, of course, then the matter becomes that much more difficult. To get a sense of which model *does* apply, we turn to the historical record.

The Supreme Court's Federalism

The Constitution's Constitutional *Decentralization*

Whether or not they *should* have done so, the Framers designed a constitutional decentralization, not a contingent one. Indeed, this choice almost certainly reflected their intuitive understanding, as expressed throughout the *Federalist Papers*, that the animating spirit of government action all too often gives evidence of the pursuit of private interest.

The Constitution itself sets out in Article I, Section 8, a collection of "enumerated" national government powers, including the traditional federative (national defense and international relations) powers and the "Power . . . to regulate Commerce with foreign Nations, among the several States, and with the Indian Tribes. . . ." The draftsmen took great pains to avoid congressional favoritism among the states in matters of trade and taxation. And they drew equally clear lines to avoid state actions in derogation of the national common market. Finally, the Tenth Amendment required that "The powers not delegated to the United States by the Constitution, nor prohibited to the States, are reserved to the States respectively, or to the people."

The Constitution's allocation of substantive powers and that provided under the Articles of Confederation are not materially different. But the Constitution also created a national executive and more robust taxing powers. It departed from the Articles' legislative decision rule (nine of thirteen states, each with one vote) to embrace majority rule with a two-house legislature, plus a potentially overridable executive veto. The Union was on firmer ground, of course, as the Constitution's supremacy clause made clear. But the overall structure of the document with respect to domestic concerns created

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

enumerated national powers "in the sense of being islands of authority in a sea of state and individual power." But with respect to federative concerns, the "powers of individual states were likewise severely limited, becoming even smaller islands of authority in a sea of national powers" (Aranson 1990). Madison expressed well the Constitution's plan of divided authority, in *Federalist* No. 14:

> The general government is not to be charged with the whole power of making and administering laws. Its jurisdiction is limited to certain enumerated objects, which concern all the members of the republic, but which are not to be attained by the separate provision of any. The subordinate governments, which can extend their care to all those other objects which can be separately provided for, will retain their due authority and activity.

The American Constitution, however, shares a problem with all other written organic documents: It is not self-enforcing. Or, as Madison again put it in *Federalist* No. 18, the Constitution provides but a "parchment barrier . . . against the encouraging spirit of power." It thus remained for the Supreme Court to lay out the metes and bounds of national and state powers, both with respect to each other and with respect to the individual citizen. This activity occurred over the intervening centuries as the result of a large number of suits, involving the federal government against a state, a state against the federal government, or either of these against an individual citizen, or vice versa. Each suit brought to the Court a claim or counterclaim, that the national or a state government had exceeded its powers absolutely, or had done so with respect to the rights or powers of another person or government entity.

The Court in such matters confronted a critical choice. It could develop bright line, doctrinal rules to define and limit state and national powers, consistent with the Constitution's words and animating spirit. Or, it could look to utilitarian purpose, to develop less a doctrine of federalism as constitutional decentralization and more a balance among competing interests, whose implications would differ from case to case, resulting in a contingent decentralization.

The Court did begin its inquiries into federalism along doctrinal lines. But these soon gave way to a consequentialist jurisprudence. The Court thereby replaced its earliest understanding of federalism as constitutional decentralization with one of federalism as contingent decentralization. The inevitable result of this shift has been the increasing centralization of the American polity, along with the decline of federalism itself.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Confining the Federal Government

The Supreme Court's problem with respect to the federal government is, was, and ever shall remain that of keeping it in its place, so that it does not assume many of the substantive powers that legitimately should remain with the states. The problem emerges in the context of several constitutional provisions. I am less concerned here with those regarding the federative powers of national defense and international relations than with those involving the commerce power. In particular: How might the Court hew out doctrinal constitutional divisions between the national and state governments?

The principal intent of the commerce power doubtless was to limit the states' abilities, through tariffs and other trade restrictions, to wage economic warfare on each other. A secondary intent was to allow Congress to enact such measures as would facilitate commerce, thereby promoting a national common market. But where, and how, would the Supreme Court draw the line on national powers?

Its earliest pronouncement on this subject, *Gibbons v. Ogden* (22 U.S. 1 [1824]), was really a case about state powers. New York had granted an exclusive franchise requiring vessels sailing to and from New Jersey to acquire and pay for the franchisee's permission to sail. A federal law arguably preempted this arrangement. Chief Justice Marshall's opinion for the Court struck down the New York law on preemption grounds. And, not surprisingly, modern commentators read the case as a brief for expansive federal powers (see, e.g., Tribe 1978, p. 232).

Richard Epstein (1987), by contrast, has discovered in the *Gibbons* opinion an important doctrinal distinction. In noting that it was within the power of the states to inspect goods bound for interstate commerce, Marshall wrote that this power is actuated *before* the goods enter the stream of commerce. Inspection laws, Marshall said, "act upon the subject before it becomes an article of foreign commerce, or of commerce among the States, and prepare it for that purpose" (22 U.S. at 203). The implication is that commerce in the contemplation of the commerce clause is a limited matter, and therefore the power of Congress to regulate interstate commerce is likewise restricted to interstate commerce itself, and does not include a power to regulate activities that occur before commerce.

Epstein traces out the rise and decline of this distinction, and with it, I would add, the rise and decline of federalism as constitutional decentralization. The distinction gained its fullest statement in *United States v. E. C. Knight* (156 U.S. 1 [1895]), which overturned a Sherman Act challenge to a merger of refined sugar manufacturers.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Prohibiting the merger, Chief Justice Fuller wrote for the Court, was not within the commerce clause powers of Congress, for the prohibition acted on *manufacture,* not on *commerce.* And "commerce succeeds to manufacture, and is not a part of it" (at 16).

This principled, doctrinal distinction held up in *Hammer v. Dagenhart* (247 U.S. 251 [1918]), but it began to unravel almost immediately thereafter. By the time of *NLRB v. Jones & Laughlin Steel Corp.* (301 U.S. 1 [1937]), Chief Justice Hughes jumped completely over the barrier that *E. C. Knight* had erected, to make the question of congressional power one of economic incidence, and not one of bright line distinctions. "It is the effect upon commerce," he wrote for the Court, and "not the source of the injury, which is the criterion" (at 32).

The complete demise of any boundary whatsoever on the commerce clause powers of Congress came in 1942, with the infamous case of *Wickard v. Filburn* (317 U.S. 111). At issue was the authority of the secretary of agriculture to fine wheat growers who consumed their entire output; that is, to fine those whose wheat never entered commerce at all. The Court found the necessary connection with interstate commerce. The possible economic effect of such wheat on interstate prices doubtless seems clear. But the doctrinal distinctions that underlie constitutional decentralization, unlike those that support contingent decentralization, should not advert to such considerations. For to impose an economic understanding on the issue of judicial line-drawing is to assert that all economic activity is bound together in a seamless web of national and increasingly international dimension. Thus, all economic activity becomes subject to congressional control.

Not surprisingly, therefore, the judicial demise of the manufacture-commerce distinction has allowed Congress to regulate virtually every aspect of enterprise, and even of enterprise confined wholly within a particular state. The commerce clause thus has become the vehicle for a plenipotentiary federal control of all aspects of private economic life, including matters of industrial organization, labor relations, worker and product safety, environmental quality, race relations, and matters involving gender.

Confining the State Governments

The Supreme Court's problems with respect to the states, of course, goes to the essential reasons for the originating allocation of commerce clause powers to the federal government. That is, the language of the Constitution plainly intended to preserve a national common market against state derogation. The Court faces problems in this

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

matter with respect to two different state activities: regulation and taxation. I consider these separately. Each subject involves exceedingly complex legal and economic questions. But a pattern emerges that parallels in many respects that found in the Court's treatment of congressional powers.

1. *Regulation.* The Court's most fully developed modern statement of its criteria for deciding whether a state regulation that is claimed to burden interstate commerce is or is not constitutionally permissible appears in *Pike v. Bruce Church, Inc.* (397 U.S. 137, 142 [1970]): "Where the statute regulates evenhandedly to effectuate a legitimate local public interest, and its effects on interstate commerce are only incidental, it will be upheld unless the burden imposed on such commerce is clearly excessive in relation to the putative local benefits."

Notwithstanding the Court's frequently announced intention to apply such balancing tests to state regulations claimed to burden interstate commerce, the Court's actual performance in this area takes a wholly different approach. In almost all modern cases, if a state statute *facially discriminates* against interstate commerce, then the Court will strike it down, unless state interests seem especially compelling or the state has an evident essential fitness to regulate.

For example, in *City of Philadelphia v. New Jersey* (437 U.S. 617 [1978]), the Court overturned a New Jersey statute that barred the importation of waste collected in other states. But compelling state interests will overcome such results. In *Maine v. Taylor* (477 U.S. 131 [1986]), for instance, the Court upheld a state law barring imported live baitfish, on the grounds that such bait might infect local gamefish with parasites not native to the state. And essential fitness to regulate comes into play in cases like the classic *Cooley v. Board of Port Wardens* (53 U.S. [12 How.] 299 [1851]), which sustained a state law requiring the use of local pilots for ships entering Philadelphia harbor, or the payment of a fine.

Where facial discrimination is absent, however, the Court almost always will sustain the challenged state regulation, unless federal preemption is present, or the state itself attempts directly or indirectly to regulate national markets. Some of these cases seem especially egregious. *Exxon Corp. v. Governor of Maryland* (437 U.S. 117 [1978]), for example, brought to the Court a challenge to a state law prohibiting petroleum producers and refiners from owning retail outlets. The statute plainly resulted from rent-seeking, entry-restricting moves by Maryland's independent gasoline retailers. But the Court sustained the law on the grounds that it applied equally to in-state and out-of-state producers and refiners. That is, it was not

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

facially discriminatory. That there might have been no domestic producers or refiners somehow escaped judicial notice.

The preemption problem emerged in *Edgar v. Mite Corp.* (457 U.S. 624 [1982]). At issue was an Illinois antitakeover statute, which reached beyond corporations chartered in that state, to impose onerous restraints on takeover activity (see, e.g., Jarrell and Bradley 1980; Jarrell 1983). The Court, in a plurality decision, overturned the statute on the grounds that the federal Williams Act preempted it.

Not so the Indiana statute challenged five years later in *CTS Corp. v. Dynamics Corp. of America* (481 U.S. 69 [1987]). There the Court sustained a law more narrowly drawn to get past its earlier holding (see, e.g., Butler 1988; Macey 1988; Fischel 1987; Langevoort 1987). Justice Scalia's concurrence puts the matter baldly: "Nothing in the Constitution says that the protection of entrenched management is any less important a 'putative local benefit' than the protection of entrenched shareholders, and I do not know what qualifies us to make that judgment" (481 U.S. at 95).

To its credit, the Court occasionally will ask if state statutes seek to circumvent national market forces. In *Hunt v. Washington State Apple Advertising Commission* (432 U.S. 333 [1977]), for example, the Court overturned a North Carolina law prohibiting the use of other states' superior agricultural grading methods in place of those developed by the Department of Agriculture. "By prohibiting Washington growers and dealers from marketing apples under their State's grades," said the Court, "the statute has a leveling effect which insidiously operates to the advantage of local apple producers" (at 351).

Brown-Forman Distillers Corp. v. New York State Liquor Authority (476 U.S. 573 [1986]) explored a similar problem. New York law required liquor prices in New York to be no higher than recent prices in other states. The law, said the Court, "regulates out-of-state transactions in violation of the Commerce Clause." It requires "a merchant to seek regulatory approval in one State before undertaking a transaction in another" and, therefore, it "directly regulates interstate commerce" (at 561–62).

2. *Taxation.* The Court's most fully developed modern statement of its criteria for deciding whether a state tax that is claimed to burden interstate commerce is or is not constitutionally permissible appears in *Complete Auto Transit, Inc. v. Brady* (530 U.S. 274 [1977]). The Court there said that it will ask whether the tax "is applied to an activity with a substantial nexus with the taxing State, is fairly apportioned, does not discriminate against interstate commerce, and is fairly related to the services provided by the State" (at 279).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Not surprisingly, the Court's jurisprudence in this area parallels in many respects that concerning state regulations that arguably burden interstate commerce. Where the tax is facially discriminatory, the Court almost always will strike it down. But where the tax is facially neutral, and notwithstanding the Court's announced intention to perform the balance suggested in *Complete Auto Transit*, the Court almost always will sustain it.

Bacchus Imports, Ltd. v. Dias (104 S.Ct. 3049), for example, brought to the Court a challenge to a Hawaii statute that partly exempted brandies and wines distilled from local produce. The Court overturned the statute on commerce clause grounds. It did likewise in *Tyler Pipe Industries, Inc. v. Washington State Department of Revenue* (483 U.S. 199 [1987]), wherein the challenged statute exempted locally made goods from a manufacturing tax if those goods would be sold in Washington and thus be subject to a wholesale tax. But the Court in *Western & Southwestern Life Ins. Co. v. State Board of Equalization* (451 U.S. 648 [1981]) sustained California's retaliatory tax on insurance premiums against firms domiciled in states whose taxes were greater than California's.

As with state regulations affecting interstate commerce, so with state taxation, and to the same effect: facial neutrality makes a tax an unlikely candidate for Supreme Court disapproval. The most egregious example is Montana's severance levy on coal, sustained in *Commonwealth Edison v. Montana* (453 U.S. 609 [1981]). Following the Arab Oil Embargo of 1973, and the subsequent increased demand for domestic coal as a source of energy for heating and power, Montana increased its severance tax on coal to as much as 30 percent (see, e.g., McLure 1982, Williams 1982).

The proceeds went to a trust fund that provided 20 percent of the state's revenues. Montana accounted for 25 percent of the nation's coal reserves and 50 percent of its reserves of low-sulfur coal. Not surprisingly, debates over the tax increase in the Montana legislature noted that other states would bear most of the burden: Montana acknowledged that nearly all of the tax would be exported, since 90 percent of the coal went to other states. But to no avail: the Court sustained the tax.

Federalism: Reasons and Rules

Our brief exploration into the economic logic of federalism, its various historical meanings, and the political and judicial forces acting upon it, suggest these conclusions. *First*, decentralization as a way to allocate responsibility for governance enjoys patent welfare-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

regarding justifications. These include cost-and-benefit responsibilities, economies of scale, and the advantages of interpolity competition. *Second,* contingent decentralization alone cannot protect against informational asymmetries and political forces leading to increased centralization. Therefore, for federalism to survive, the courts must be willing to apply the original understanding of the American arrangement as *constitutional* decentralization. *Third,* the Supreme Court has turned aside from this task, with two inevitable results: the federal government has gained power at the expense of the states and of the welfare of citizens that constitutional decentralization originally contemplated; and paradoxically, the states have gained power over the national common market, with similar unwanted consequences for that market's preservation. It remains for us to close three interlocking circles with respect to these conclusions.

The Reason of Rules

First, if we advert to the utilitarian, welfare-regarding goals and justifications for *any* form of decentralization, our exploration shows that the decentralization that makes the achievement of those goals possible must rely for its existence on ignoring those goals entirely in particular cases. That is, the reason of the doctrinal rules of federalism lies in the paradoxical claim that the decentralization will fail, and therefore the utilitarian consequences of decentralization will disappear, if we invoke those consequences alone in our defenses of decentralization.

The demise of federalism as contingent decentralization, and tautologically the demise of federalism as constitutional decentralization, begins with the meta-decision to choose sides in particular cases solely by considering consequences instead of rules and rights. Balancing tests in place of doctrinal constitutional argument provide the acid that eats away at the girders of federalism, leaving the states not more than *functus officio,* except as they can here and there sequester the ability to erode market forces that the constitutional arrangement so obviously sought to protect. One can do no more, in this respect, than to quote Hayek's (1973, p. 61) famous passage about the defense of freedom, and draw the straightforward analogy about the defense of federalism:

> The preservation of a free system is so difficult precisely because it requires a constant rejection of measures which appear to be required to secure particular results, on no stronger grounds than that they conflict with a general rule, and frequently without our knowing what will be the costs of not observing the rule in the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> particular instance. A successful defense of freedom must therefore be dogmatic and make no concessions to expediency, even where it is not possible to show that, besides the known beneficial effects, some particular harmful result would also follow from its infringement. Freedom will prevail only if it is accepted as a general principle, whose application to particular instances requires no justification.

The Limits of Rules

But are there limits to doctrine? After all, applying such an approach to a jurisprudence of federalism must seem a terrifying prospect. As Justice Jackson noted in a different connection, in his dissent in *Terminiello v. Chicago* (337 U.S. 1, 37 [1949]): "If the Court does not temper its doctrinaire logic with a little practical wisdom, it will convert the constitutional Bill of Rights into a suicide pact." That possibility raises two distinct problems.

First, a doctrinal adherence to federalism as constitutional decentralization actually might subvert the achievement of its particular utilitarian goals. For example, air or water pollution might flow across state lines, and the optimal scale of some government services might exceed the output of particular, and especially smaller states.

I am not greatly concerned about such prospects. The question is *not* whether we should subvert the constitutional guarantees of federalism, to force the states to address such problems or to have the federal government do the job for them. There may be other constitutional grounds for national action. More important, both federalism in general, and the Constitution in particular, embrace the possibility of joint action among sets of states to overcome these problems. Where mutually beneficial gains from joint action are possible, they can emerge voluntarily, as in any market relation.

Surely, the states have not done enough using such tools. In any potential contract, the risk is inevitable that jockeying for position in dividing the surplus of joint action will undermine agreement. But the inadequacy of present interstate cooperation also gives evidence that the tools of agreement have grown rusty precisely because of federal preemption. The quality of states' cooperative actions, therefore, may remain a function of the national government's willingness or unwillingness to stand aside, to allow learning to occur. That course, I believe, holds better prospects than does the present use of national legislation, and even that legislation regarding exactly such questions as pollution control, for mutual plunder (see Ackerman and Hassler 1981).

Second, there may remain a residuum of issues on which people hold intense preferences, and where progress might not occur with-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

out a frank overturning of federalism's constitutional predicates. The underlying political logic of such situations is thus. People within or beyond the borders of a particular state want that state's political or judicial branches to take actions that a majority of its citizens or agents do not prefer. But the demanders know that agents located in national political or judicial branches are willing to take such actions. Federalism, however, allocates responsibility for such actions to the states. Federalism thereby protects the preferences of (some of) those within the state, to the detriment of others within or without the state.

The central most important issue that this pattern affects is race. Riker (1964, p. 155), writing over a quarter of a century ago, emphasized both this aspect of federalism and his own results-oriented approach to it:

> If one approves the goals and values of the privileged minority, one should approve the federalism. Thus, if in the United States one approves of Southern white racists, then one should approve of American federalism. If, on the other hand, one disapproves of the values of the privileged minority, one should disapprove of federalism. Thus, if in the United States one disapproves of racism, one should disapprove of federalism.

Riker's consequentialist view leaves us with two alternatives: Enjoy the virtues of federalism along with the vices of institutionalized racism; or, forego the virtues of federalism and achieve the virtues of racial justice. If the choice is as stated, I am no better than is anyone else at deciding, though I am fairly certain that I would forego federalism, with the hope that some aspects of it could be resurrected after justice prevails.

But even in such circumstances, the facts are seldom so clear. First, the citizen's best protection may be the ability to avoid the transgressions of racism by voting with his feet. Foregoing federalism eventually might entail the loss of such protection. Second, the nationalized civil rights revolution allowed state officials perpetually to avoid the responsibility of making hard choices. That, after all, was the implication of congressionally and judicially mandated desegregation orders.

Yet, they might have had to make such choices had they been left to their own devices, and had they been made to face political developments within and without their states. Much of the civil rights revolution, for example, grew out of the dominance in the Senate of Northern liberal Democrats, following the presidential election of 1964 (Aranson 1981, chap. 4). Before that time the Southern states economically had been kept afloat by a stream of federal

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

spending that relied on the dominance in Congress of Southern Democrats, who held important committee leadership positions by virtue of their seniority. The elections of 1964 jeopardized that pattern of control and expenditures, and the Southern states soon would have had to relent on issues of race, in exchange for continuing federal spending in the region.

We will never know how much better it might have been for peaceful change and for the quality of state leadership had the Southern states instituted their own changes. But the matter deserves some thought, especially because of the loss of a robust federalism that the conflict's national, constitutional resolution ultimately imposed on us.

Nor is it clear that the loss of federalism's protections brought unalloyed benefits in other policy areas. The value of some of these protections we now can acknowledge in hindsight and in light of the present Court's occasional, though meager and ultimately misguided attempts to resurrect the elements of constitutional decentralization.

Federalism can be a neutral sword in its defense of ideological positions *within* states. Certainly, the religious right gained the advantage on the grounds of federalism when the Court upheld Missouri's more restrictive regulation of abortions, in *Webster v. Reproductive Health Services* (109 S.Ct. 3040 [1989]). But the Court in *Bowen v. American Hospital Ass'n* (476 U.S. 610 [1986]) likewise turned aside on grounds of federalism the Reagan administration's attempt to regulate, and ultimately to eliminate, the practice of withholding life-saving medical procedures from impaired newborn children.

Even Governor Michael Dukakis, though ultimately unsuccessful, could advance reasons of federalism in support of his preventing the training of Bay State National Guard personnel in Latin America. The ultimate irony doubtless will occur on the day that Governor Douglas Wilder's attorney general argues before the Supreme Court that the federal government has no business telling the governor of Virginia how to conduct his state's affairs.

The Limits of Reason

The final circle to close reflects the worrisome limits of reason in defending or attacking ancient practices such as constitutional decentralization. We cannot find the ultimate defense of such practices in particular results, for those results accept a static reality, with no possibility for real institutional change. An argument in favor of national action, and against constitutional decentralization, for example, may take this form: Certain questions are sufficiently com-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

plex and conflicted that they require a more progressive, sophisticated, responsive, and courageous legislature for their resolution. Concerning such matters, the federal House and Senate are thus to be preferred to their counterparts in the states.

The problem with this claim, aside from its ultimate unbelievability in the light of FSLIC and related scandals, is that the quality of representation is not exogenous to the problems to be resolved. If federalism as constitutional decentralization perishes, and if the states remain only with low-level administrative problems to resolve, then citizens will give state governments little heed. Not surprisingly, far more Americans know the names of their federal senators and representatives than those of their statehouse counterparts. If state government matters little, then elections of state officials will matter less. An improvement in the quality of state governments thus may require state political leaders to engage just those issues with which, by hypothesis, they remain not yet ready to struggle.

This observation brings us full circle to the question with which we began. The Framers' meaning of federalism was constitutional decentralization, and we have traced out how that meaning and its implications have been lost to us. But before the Framers' federalism, the idea of federalism embraced an older concept of covenant, as among sovereigns. Covenant preserved the dignity and worth of each sovereign, and as applied by analogy to the individual human being, it preserved first self-actualization and essential "humanity" and then republican virtue.

But these valuable aspects of human existence did require polities to be small enough to allow meaningful, informed, and efficacious participation in political life. Those adjectives measure goals whose achievement reflects one characteristic above all else: the size of the polity. A vote in a national election shares nothing but symbolic worth. A vote in New Hampshire, or Vermont, or Alaska, by contrast, begins to have some public consequence, however small, beyond the voter's symbolic satisfaction.

Can citizens care about the public consequences of their actions if no such consequences prevail? I believe that that is not more likely than that they can care about (state) legislators whose functions are grossly limited by federalism's demise. Perhaps the devolution of such responsibility to the individual citizen through federalism's reinvigoration will produce no such public-regardedness. The very concept might have been lost forever, along with the notion of republican virtue itself, as a consequence of two centuries of centralization. That is, our contemporaneous reason may have no place for such concepts in calculating the virtues of any measure of decentraliza-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tion. But we shall never know how robust these concepts might be until we put them to the test. Meanwhile, we have Montesquieu's promise that a republic, "if it be large, . . . is ruined by an internal imperfection."

References

Ackerman, Bruce A., and Hassler, William T. *Clean Coal/Dirty Air*. New Haven: Yale University Press, 1981.

Aranson, Peter H. *American Government: Strategy and Choice*. Boston: Little, Brown, 1981.

Aranson, Peter H. "The European Economic Community: Lessons from America." Paper prepared for L'Université d'Été de la Nouvelle Economie, Aix-en-Provence, 1989.

Aranson, Peter H. "Rational Ignorance in Politics, Economics, and Law." *Le Journal des Economistes et des Etudes Humaines* 1 (1989/1990): 25–42.

Aranson, Peter H. "Federalism: Doctrine against Balance." Paper prepared for a Liberty Fund Conference on Freedom and Federalism, Crystal City, Va., 19–22 April 1990.

Aranson, Peter H., and Ordeshook, Peter C. "Public Interest, Private Interest, and the Democratic Polity." In *The Democratic State*, pp. 86–177. Edited by Roger Benjamin and Stephen Elkin. Lawrence: The University Press of Kansas, 1985.

Brennan, Geoffrey, and Buchanan, James M. *The Reason of Rules: Constitutional Political Economy*. New York: Cambridge University Press, 1985.

Butler, Henry N. "Corporation-Specific Anti-Takeover Statutes and the Market for Corporate Control." *Wisconsin Law Review* 1988 (1988): 365–83.

Diamond, Martin. "On the Relationship of Federalism and Decentralization." In *Cooperation and Conflict: Readings in American Federalism*, pp. 72–82. Edited by Daniel Elazar, R. Bruce Carroll, E. Lester Levine, and Douglas St. Angelo. Itasca, Illinois: F. E. Peacock Publishers, 1969.

Elazar, Daniel, "Foreword." In Vincent Ostrom, *The Political Theory of the Compound Republic: Designing the American Experiment*, 2nd ed. rev. 1987.

Epstein, Richard A. "The Proper Scope of the Commerce Power." *Virginia Law Review* 73 (November 1987): 1387–1455.

Farrand, Max, ed. *The Records of the Federal Convention of 1787*. New Haven: Yale University Press, 1937.

Fischel, Daniel R. "From Mite to CTS: State Anti-Takeover Statutes, the Williams Act, the Commerce Clause, and Insider Trading." *Supreme Court Economic Review* (1987): 47–95.

Hayek, Friedrich A. "The Use of Knowledge in Society." *American Economic Review* 35 (September 1945): 519–30.

Hayek, Friedrich A. *Law, Legislation, and Liberty*, Vol. 1. Chicago: University of Chicago Press, 1973.

Jarrell, Gregg A. "State Anti-Takeover Laws and the Efficient Allocation of Corporate Control: An Economic Analysis of *Edgar v. Mite Corp. Supreme Court Economic Review* 2 (1983): 111–29.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Jarrell, Gregg A., and Bradley, Michael. "The Economic Effects of Federal and State Regulation of Cash Tender Offers." *Journal of Law and Economics* 23 (October 1980): 371–407.

Jensen, Merrill. *The Articles of Confederation: An Interpretation of the Social-Constitutional History of the American Revolution, 1774–1781.* Madison: University of Wisconsin Press, 1940.

Kitch, Edmund. "Regulation and the American Common Market." In *Regulation, Federalism, and Interstate Commerce,* pp. 9–55. Edited by A. Dan Tarlock. Cambridge, Mass.: Oelgeschlager, Gunn & Hain, 1981.

Krueger, Anne O. "The Political Economy of the Rent-Seeking Society." *American Economic Review* 64 (June 1974): 291–303.

Langevoort, Donald C. "The Supreme Court and the Politics of Corporate Takeovers: A Comment on *CTS Corp. v. Dynamics Corp. of America.*" *Harvard Law Review* 101 (November 1987): 96–118.

Macey, Jonathan R. "State Anti-Takeover Legislation and the National Economy." *Wisconsin Law Review* 1988 (1988): 467–90.

Macey, Jonathan R. "Federal Deference to Local Regulators and the Economic Theory of Regulation: Towards a Franchise Theory of Federalism." Paper prepared for a Liberty Fund Conference on Freedom and Federalism, Crystal City, Va., 19–22 April 1990.

McLure, Charles E. "Incidence Analysis and the Supreme Court: An Examination of Four Cases from the 1980 Term." *Supreme Court Economic Review* 1 (1982): 69–112.

Montesquieu, Charles Secondat, Baron de. *The Spirit of the Laws.* Trans. T. Nugent. New York: Hafner, 1949.

Oates, Wallace E. "An Economist's Perspective on Fiscal Federalism." In *The Political Economy of Fiscal Federalism,* chap. 1. Edited by Wallace E. Oates. Lexington, Mass.: Lexington Books, 1977.

Ostrom, Vincent. "Courts and Collectivities." Paper prepared for a Conference on Public Choice and the Judiciary, Utah State University, 1989.

Peltzman, Sam. "Toward a More General Theory of Regulation." *Journal of Law and Economics* 19 (August 1976): 211–40.

Riker, William H. *Federalism: Origin, Operation, Significance.* Boston: Little, Brown, 1964.

Riker, William H. *The Development of American Federalism.* Boston: Kluwer Academic Publishers, 1987.

Riker, William H. "The Lessons of 1787." Unpublished manuscript, University of Rochester, n.d.

Stigler, George J. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (Spring 1971): 3–21.

Tiebout, Charles. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (1956): 416–24.

Tribe, Lawrence. *American Constitutional Law.* Mineola, N.Y.: Foundation Press, 1978.

Tullock, Gordon. "The Welfare Costs of Tariffs, Monopoly, and Theft." *Western Economic Journal* 5 (June 1967): 224–32.

Weingast, Barry R.; Shepsle, Kenneth A.; and Johnsen, Christopher. "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics." *Journal of Political Economy* 89 (August 1981): 642–64.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Williams, Steven F. "Severance Taxes and Federalism: The Role of the Supreme Court in Preserving a National Common Market for Energy Supplies." *University of Colorado Law Review* 53 (Winter 1982): 281–314.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A Double Security: Federalism as Competition

Aaron Wildavsky

> A double security arises to the rights of the people. The different governments will control each other, at the same time that each will be controlled by itself.
>
> —*Federalist* No. 51

Introduction

The triumph of America's regime values of democracy and capitalism in Eastern Europe would appear to justify a brief moment of self-congratulation. In a thought-provoking essay, Francis Fukuyama (1989) has memorialized "the end of history," not, of course, of events but of differences over which type of economic institution is most productive (i.e., markets) and what kind of political process (namely, democracy) best promotes human welfare. I think not. No one seems to have noticed the effort to use democracy against capitalism in the name of equality.

The existing poverty of organizational or cultural forms has boxed human choice into a dichotomous straitjacket: On one side is hierarchical collectivism, as in feudal monarchy or Soviet communism; on the other, competitive individualism, as in American capitalism.[1] Since there are only two possible places for human beings to end up, if they reject one (say Soviet-style hierarchies), they must embrace the other (some sort of capitalism). Not so. Egalitarianism—a life devoted both to community and to individual expression; a life in which the guiding norm is greater equality of condition and in which

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Political Science and Public Policy at the University of California at Berkeley. "Federalism as Competition" is the title of the first chapter in Anton (1989).

[1]For an extended discussion demonstrating the dualism of the masters of social thought from Montesquieu to Merton, see Thompson, Ellis, and Wildavsky (1990).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the gradations of power are suspect of depriving citizens of participation in the public sphere—remains a lively option, nowhere more so than in the United States.[2]

Interpretations of American history have focused on the defeat of hierarchy by individualism. The exceptional features of American political life—the absence of strong socialist parties or nationalized industries or nearly as large a welfare state apparatus as in Europe—have been traced, following Louis Hartz's influence, to the absence of feudalism, that is, to the lack of a hereditary hierarchy.[3] While this interpretation is true and important, it is not all the truth—and I increasingly believe it is not the main truth.

True enough, from the revolutionary period on, the principle of hierarchy has grown steadily weaker. It is estimated that a third of those who identified themselves as Tories (hierarchists to us) fled to Canada. The only two avowedly hierarchical major parties, the Federalists and the Whigs, were short-lived, unable to elect presidents of their own persuasion (after the two Adamses died) before the Civil War. To say this without saying more, however, is grossly to distort American history because the statement leaves out the crucial part played by egalitarianism.

What was exceptional about the United States of America was not only the weakness of hierarchy but its defeat by an unusual coalition of individualists and egalitarians. The origins of this coalition[4] lay, first, in the mixture of these two cultures in the immigration of Puritans,[5] and, second, in their common opposition to that distant (and, they felt, oppressive) hierarchy, King George the Third. Egalitarian influence was strong among the anti-Federalists, was stronger still among Jeffersonians and Jacksonians, and was dominant among Abolitionists (Friedman 1982, Ellis and Wildavsky 1990) and Populists (Malecha 1988). Egalitarian influence is now strong among Democratic Party activists (Jennings and Miller 1986, Polsby and Wildavsky 1988); environmentalists (Dake and Wildavsky 1990, Douglas

[2]On egalitarianism, see Douglas (1970), Douglas and Wildavsky (1982), and Ellis and Wildavsky (1989).

[3]Hartz's *The Liberal Tradition in America* (1955) chronicles the supremacy of John Locke's liberalism, harking back to the continental use of "liberal" as "individualistic."

[4]I believe that this same coalition is responsible for the rise of capitalism (see Webber and Wildavsky 1986).

[5]Thompson, Ellis, and Wildavsky (1990) reanalyze Max Weber's splendid *The Protestant Ethic and the Spirit of Capitalism* to show that he conflated rather than separated the quite different opposition of individualists and egalitarians to the hierarchy of their day.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and Wildavsky 1982); and major media journalists (Lichter, Rothman, and Lichter 1986; Wildavsky 1987).

From colonial times until the mid-1960s, the individualist-egalitarian alliance was based on the belief that equality of opportunity, without the unfair intervention of the central government, was compatible with sufficient equality of condition to support liberty. In short, American egalitarians located the source of inequality in the central government. The difference is that their ideological successors locate inequality in corporate capitalism. Instead of an alliance of egalitarians and individualists against hierarchists, egalitarians are on the attack against both, with emphasis on opposition to individualism. So what? Does this radical change matter? And why is it being discussed in a paper on federalism?

The worst mistake one can make in interpreting our political history, in my opinion, is to merge individualists and egalitarians as people who oppose authority (Wildavsky 1988b). Though individualists prefer to substitute self-regulation for authority, they wish to minimize, not to overturn it. By contrast, egalitarians reject authority as inherently inegalitarian. Where individualists are technologically optimistic, believing that risk is opportunity, egalitarians (after the Industrial Revolution) came to see large-scale industry as a producer of inequality and hence morally polluted. In the view of egalitarians, therefore, industry should be regulated extensively to reduce its harms. And there is the rub.

The rise of egalitarianism matters to students of federalism because egalitarians oppose competition when they see it as a source of inequality; yet there can be no competition without creating inequality. Instead of espousing, as they used to do, that equal opportunity and equal condition are compatible in America, egalitarians now contend that there can be no genuine opportunity without much greater equality of condition. Whereas they once legitimated political procedures without requiring agreement on policy, they now demand agreement on egalitarian measures as a precondition of granting legitimacy to the political system.

Uniformity is antithetical to federalism. A belief in uniformity, for instance, could be manifested by strong central rule and modified by delegating authority to geographic units. This delegation, though it belongs to a unitary conception of government, is often called decentralization. Noncentralization, by contrast, refers to independent centers of power in geographic areas that do (and are expected to) differentiate themselves. Though there may be some centralization in a federal system, and there might be decentralization (i.e., delegation), there must be noncentralization.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A belief in equality, not only of opportunity but of outcome, would be hostile to noncentralization, for then there could be no substantial differences among states. The existence of states free to disagree with one another and with the central government, however, inevitably leads to differentiation. Yet states must differ with each other and with the national government if they are to do more than obey central directives. Were there to be a change in values toward equality of condition, therefore, the political culture that undergirds federalism would fall apart. For if there is anything essential to a vibrant federal system, it is competition.

The Three Progressives

Three of the great movements of modernity—democracy, science, and capitalism—are based on competition. Democracy is centrally concerned with competition for office, including the auxiliary liberties necessary to make it work—information, speech, association, and voting. Those who have suffered most from dictatorship, we all observe, place great value on free elections. The activity we call science is well conceived as competition over ideas. Knowledge is sought through what Karl Popper (1962) calls "conjectures and refutations" (i.e., competition). The theory and the practice of economic growth is based on competition. The idea is that, in national and international markets, investors seek the highest return, thus reinforcing the highest value ventures and weakening the worst. When competition is absent, as in command economies, growth declines. For example, a study of Polish planning shows that economic action by issuing commands creates negative selection; with unerring precision and regularity, the least economically worthwhile projects are chosen (Clark and Wildavsky 1990).

When we want knowledge, we do not set up a hierarchy whose purpose is to figure out the best way to learn new things; on the contrary, we rely on a competitive process in which numerous people, connected by common interests, strive for excellence. When we want liberty, we do not establish an organization to make us free; instead, we rely on political parties to compete for popular favor. When we want economic growth, we do not set up a command and control system, because we know that would be counterproductive. Rather, we rely on bidding and bargaining among independent and interdependent economic agents to make resources grow. Why, then, when we wish to arrive at a division of labor between the national and the state governments, do we seek a central, cerebral solution, as if there could be some set of principles to be accepted by (or

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

imposed on) everyone that would tell us what should be done by whom? If we think so well of competition that we enthrone it in democracy, science, and economics, why should we not tackle the problem of federal structure in the same way? What would a federal system look like if it were organized by rules meant to produce a division of labor that would reflect the results of competition?

A Prelude to Competition

Parallel play (a space for states, a space for national government, and a space for both of them together) was once the neat theory of the American federal system. Under the prevailing image of the layer cake (a platonic arrangement if ever there was one), the parts never penetrated except by benefit of constitutional clergy. When Morton Grodzins (1966) observed that this parallel plane was really a triangle, a marble cake instead of a layer cake, there was confusion but not consternation. As Daniel Elazar (1962) demonstrated, cooperative federalism was the norm virtually from the outset. The good feeling remained that American pragmatism had apparently triumphed over foreign theorizing.

The question of what exactly to call this compound of national and state rule was, for a time, superseded by calling it "good." So long as a few bands of dark were still visible against the white, there was sufficient resolution to say it was something. We could have our federal cake (color me marble!) and eat it too. But as government grew from the mid-1960s, bands of marble gave way to veins going every which way, so crisscrossed that no one could say what was up or down or who was (simultaneously?) on top or bottom. Big government (large numbers of large programs taking a larger proportion of national income) had made the division of powers among governments unrecognizable; it was hard to tell one level from another (say, with the federal government supporting local libraries) or to discern any difference in principle among the levels.

Under a unitary regime, states and localities carry out national instructions; the problem is how to improve their obedience. In a federal regime, states and localities are disobedient. The operational meaning of federalism is found in the degree to which the constituent units disagree about what should be done, who should do it, and how it should be carried out. In a word, federalism is about competition and conflict. Federalism is also about cooperation, that is, the terms and conditions under which competition is limited. A federal regime, therefore, cannot be coordinated any more than it can be controlled or coerced. Federalism requires mutuality instead of hierarchy, mul-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tiple rather than single causation, and a sharing instead of a monopoly of power.

If there is a federal principle, it cannot be limited to relations between national and state governments. If it is good for power to be divided and shared, that principle must also prevail in relationships among states and their cities, counties, and special districts. Let us call this principle—under most conditions, a larger number of smaller units will deliver services better than a smaller number of larger units—the federal bias. It is not an unalterable bias; circumstances do alter cases. It is, however, a bias. The burden of proof, for those who share this bias toward federalism, must be on people who propose that bigger and fewer are almost always better (Wildavsky 1976).

If federalism is institutionalized competition among governments, increasing rather than limiting a citizen's choice among service providers ought to be its contemporary key. Citizens will be better served by our asking how to improve their choices rather than by our worrying over which level of big government should monopolize a service. Students of federalism should look not only at the balance among levels of government but, more importantly today, at the relative proportions of the public and private sectors so that competition among governments may be enhanced. When a citizen's choice characterizes federalism, which is as it should be if we want government to adjust to individual taste rather than adjusting the people to their government, the programmatic products should differ state by state.

However, the basic proposition of bureaucratic government, which runs across state and national lines and is fortified by conventional administrative wisdom, is monopolist: Governments should control the allocation of services to citizens within their jurisdictions without competition. If governments had their way, each level would seek to limit competition by getting a higher level to guarantee it. What you see is what you get: The dinosaur solution operates so that programs multiply without simultaneously increasing the intelligence of those who design and administer the programs.

In private industry the theory is competitive but the practice sometimes is not. Following William Niskanen (1971), we may stipulate that, if it could get its way, industry would prefer competition among its suppliers and monopolistic access to its consumers. That is why it is necessary to bring external force to bear to ensure competition. The same sort of pressure must be brought to bear on government if it is to give up its monopolies. No one likes competition (after all,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

you can lose) except the consumer whose life is made predictable by the uncertainties imposed on others.

For the citizen, then, overlap, duplication, and redundancy (synonyms, in this context, for competition) are not wasteful but wonderful. They enhance reliability of supply; if one supplier will not, another will. Competition cuts down information costs; products may be distinguished by prices, exchanged, or given diversity; their mixes may be compared at the margins. Redundancy results in suppliers working hard to guess what citizen consumers want rather than the reverse. Competition is the better mouse trap that has a variety of units competing to beat a path to the citizen's door.

Anyone who has to deal with governmental agencies, or who wants to apply for grants, knows that it is useful to have a multiplicity of points of access and largess. It is then possible to play them off against each other or simply to go to the most receptive place. Individuals can thus sort themselves out among agencies with different rules, objectives, and personal predilections. Similarly, grievances encountered in one place may be redressed in another. There is no a priori mode of action that fits everyone; instead of fitting citizens into a mold, that is, requiring that they be uniform, it is better to demand that agencies be diverse.

Who is to say what is the appropriate scale of activities? Outside a limited area, including perhaps water and sewers, many unsubstantiated claims are made for economies of scale. The division of labor facilitated by large size may suit certain specializations but is unlikely to result in more cops on the beat or teachers in the classroom or case workers in welfare.

The major movement in education in the United States after World War II, for instance, has been the consolidation of rural school districts. Every study shows that the larger the size, the worse the scholastic or social performance—no matter what measure is used. One understands that consolidation may be useful to state and national associations or supervisors or may reduce the number subject to governmental oversight, but it is a bad buy for citizens subject to it (see Niskanen 1990).

A competitive approach will lead to better, if not bigger, state government. Consider, for example, the quandary of state governments in dealing with pensions and disability payments for their employees. Given the normal operation of the political process, it is all too tempting to hide the true costs of pension provisions by pushing them into the future when another generation will have to pay. The future apparently is now, and the result is huge "unanticipated" costs. The direct way of dealing with this phenomenon is to

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

set up special committees to monitor retirement provisions. Several states have done just this. But, after the first flush of enthusiasm wears off, a return to business as usual may be expected. Contracting out, by contrast, requires no constant surveillance. Whatever faults the agreement has, other people—the service providers—are responsible for fringe benefits. All the financial costs are on top, out in the open. The relationships are easier to monitor, thus leading to more effective and responsive government.

Nevertheless, if conditions stay the same, monopoly will grow. Indeed, feeling their oats, monopolistic bureaus will not be able to understand why reason and rationality do not dictate that certain classes of activities be provided exclusively through government.

What, then, should be done if one desires a federal system that stresses variety, diversity, and competition? As usual, negative knowledge comes earlier than positive knowledge—big government is antithetical to federalism because it preempts resources for the national government, and because it causes virtually every activity to mix the levels of government. Variety in programs suffers as the feds seek as much conformity as they can get. The point I wish to pursue about positive knowledge is that we do not need what we should not have. A cognitive approach through a formula for demarcation suggests that there can be a general theory to which government activity should be subject. Only citizens, through repeated iterations, can work that theory out for themselves. Let us turn instead to setting up those conditions for social interaction that are most likely to produce outcomes that conform to citizens' preferences.

The Federal Government as the Fifty-First State

The first thing is to reduce the federal tax to a considerably lower proportion of national income. This reduction is beginning to happen. It is permissive, and it allows states room into which they can expand if their citizens are willing to pay. The thing is not to prejudge the issue; when you do not know what to do, why do anything? Let the states pick up what they wish; let the feds give up what they can. And may the best government win or even, dare I say so, lose weight.

Second, we should think of the federal government as the fifty-first state, that is, as a source of diversity rather than uniformity. Why, the usual question goes, must the feds step in where states fear to tread? Because, the answer always is, a uniform response is required. Since it is diversity that we desire, however, it would be better to ask whether citizen consumers have sufficient variety from which to

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

choose and, if not, to add rather than subtract alternatives from the menu of public policies.

Is "interaction" a synonym for "inactivity"? Not necessarily. The federal government can act as the fifty-first state by encouraging variety through competition whenever it judges that another alternative ought to be offered to the people directly (such as competitive medical plans) or to the states (better service delivery).

Suppose no government wants to buy some good or service that some people deem essential? Presumably the program has failed an essential political test in a democracy. Suppose some states offer certain services and others do not; one or the other will be disadvantaged. Exactly. That is what should happen to achieve variety and diversity. Maybe what is really meant is that some states will end up with all the bads—the poorest people, the most noxious industries—and none of the goods. Then I suggest the people fire their incompetent politicians and hire competent ones.

Imagine states that keep out undesirable people and industries. Other states may then pick up their jobs and their people. A policy of ecology without people, or of people without industry, or of industry without either, makes no sense. Let people sort themselves out. No telling what they will do. Why, they might even enjoy themselves in ways not yet discovered.

Division by Cognition or by Interaction

Where is the philosopher's stone of federalism, the demarcation principle between central and state functions that would rationalize a division of labor within the American federal system? Nowhere. Every effort to specify what is local and what is national, I believe, will founder on these facts of life: Every national activity has its local aspects and every local activity has a national perspective. Any two incompatible objectives may be reconciled by a third to which they contribute, and any two that are complementary may be made contradictory in reference to their contributions to a third. It follows that no criterion of choice can be consistent in regard to all the legitimate political perspectives that may be brought to bear on them.[6]

If there is to be a division of policies among levels of government, there must be a mechanism for doing it. Yet there is no universally approved formula for a division of functions that commands assent. A cognitive or planning approach, under which the correct decisions flow from agreed premises, therefore, is not feasible. Instead, I rec-

[6]The extensive and inconclusive literature on the idea of a demarcation criterion separating science from nonscience should serve as a warning to enthusiasts.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ommend an interactive approach designed not only to implement but to discover what works best. The greater the number and variety of organizations for providing goods and services from government, which constitutes a competitive, citizen-centered approach, the more flexible the system will be to respond to taxpayer-voter preferences.

The people do not have to postulate either the optimal size of producing units or the optimal pattern of spatial structure. Rather these would be seen as the outcome of citizens, states, and localities getting the best deal they can and correcting errors as they go along—provided there is an effective financial constraint so that collusion (receiving benefits without paying taxes) is ruled out. Because one cannot predict policy outcomes (if we could, we could also achieve the optimum by planning) or posit the most appropriate governmental structures, the uncertainty may prove unattractive.

Some of us have assumed that nothing good occurs unless some bureaucracy makes it happen. Well, we have been down this road and it is a dead end. A competitive approach may make former monopolists more nervous but it should also leave consumers of services more confident. If the question is who should bear the costs of change, citizens as sovereigns should pass those costs on to those who supposedly serve the public.

As Mark Sproule-Jones (1979) concludes,

> The key requirement for institutional reform is to separate the physical production or supply of public goods from the jurisdiction that articulates and aggregates the demands for such services. The economic benefits from reorganization, in terms of internalizing jurisdictional spillovers and exploiting economics of large scale production, may then be harmonized with the advantages—both in economics and in responsiveness—of exposing urban bureaucracies to the threat of potential rivals in production. This reform also offers the potential of enlisting citizen participation in the implementation of public policies from which they have been largely excluded.
>
> The political feasibility of such an institutional reform rests on the relative power of public officials to maintain the monopoly status of their organizations. It would be sanguine to believe that such a reform could be instituted without a struggle.

The desirability of reducing monopoly statuses explains why it is important to spread the federal principle to lower levels. Federalism is misconstrued if it is conceived to be about monolithic structures—national and state—clashing by day, cooperating by night. If competition is desirable among states, why is it not desirable among cities, counties, and special districts? As Buchanan and Brennan (1980, p. 186) put it so well, "restatement of the familiar from one setting becomes a challenge to orthodoxy in another." Vincent Ostrom (1990)

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

states the competitive federalism project even more grandly, "we would have the emergence of human societies that no longer depend upon a unity of power to achieve coherence in patterns in human societies. Such an idea is of radical proportions." This vision of variety might not be so bad: The federal government learns to perform a narrower range of tasks better. States learn to live with widely varying styles of life. People vote by moving to a place where residents carry on the kind of life they would like to live. Perhaps that kind of active personal choice among governmental lifestyles may be as good as we can get.

Requirements of Competitive Federalism

Once we give up on the idea of creating a set of criteria that will tell us which governmental activities should be determined at the national level, which by the states, and which shared, we are still left with determining the structural principles that would facilitate a competitive federal system. Because we are talking about politics not economics, voters and parties not entrepreneurs, votes not shares, then analogies to markets have to be made with care. Nor is there the agreement on bedrock requirements for flourishing markets that one might wish. Were there such an agreement and understanding, we would not hear nonsense nowadays about so-called "halfway houses" between capitalism and socialism (private markets and Soviet-style command economies). Loose talk in the past about how a certain size government (as a proportion of GNP) would destroy liberty has been refuted by events in which countries such as Sweden and the Netherlands have spent above 60 percent of their GNP while not only being democratic but also experiencing considerable economic growth. Criticism of the element of competition that I put forward is welcome; so are substitutes.

Mechanisms of Selection

Since competition may be viewed as a process of evolutionary adaptation, I shall begin with the mechanisms of selection. Citizens vote in at least three ways to manifest their preferences about "who should do what and how much" in our federal system. First, they vote with their feet by moving from one state's jurisdiction to another, much as East Germans expressed their preference by moving en masse to West Germany. Motivation to move for a single reason only, such as the combination of taxes and services, is unlikely; other factors, such as climate, housing, family, and job, may figure importantly. Nevertheless, government may be a consideration, a factor

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

amplified by the propensity of Americans to move in large proportion and often. Presumably, two things will happen: Governments will try to make themselves more attractive to people, and people of diverse preferences should be able to find niches that come closer to suiting them.

The second way for selection to be effective is there must be an incentive to diversify. One such incentive is the necessity of attracting and retaining people. Another is the desire of rival political parties to obtain and retain control of government by competing for votes. Those people who are interested in competitive federalism should want to make it as easy as possible for new parties to organize so as to preserve and intensify competition. For instance, these people should support legislation to remove restrictions on party organization, including a party's ability to endorse whom it pleases, to restrict primaries to citizens registered with that party, and to spend however much it pleases in support of its candidates. Because parties depend on developing loyalty, they should have every opportunity to place their brand on candidates so as simultaneously to sharpen competition among parties and to simplify choice for voters. Were it deemed advisable to assist weaker and poorer parties to compete in elections, state-supported floors would be preferable to ceilings.

For the same sort of reasons, no effort should be spared in making it easier for citizens to register, which is often the greatest obstacle to voting. Aside from the advantage that higher rates of voting give to the legitimation of the overall political system,[7] we see that the larger the turnout and the higher the proportion of those who engage in this competition, then the harder the political parties have to work at being competitive and the more certain (according to the law of large numbers) is the process of selection.

A third mode of selection is a combination of referenda and polls of public opinion. Referenda like Proposition 13 in California send signals about voter preferences far and wide. Opinion polls, when repeated often enough, when independent from each other, and when using varied methods, provide signals for selection. These signals are how we know, for instance, that citizens, while desiring a lower deficit, are strongly opposed to increased income taxes (White and Wildavsky 1990). Of course, a courageous (or foolhardy) party could try to alter existing preferences, but it would know what the bulk of the public thought.

[7]The United states ranks high in the proportion of *registered* citizens who vote but low on the proportions of *total* citizens who exercise the franchise.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Effects of Information on Selection

Selection depends on information. And competition is the best disseminator there is. Without competition, for instance, cost information could not be obtained. If the price of, say, mailing a letter is what anyone will charge for delivery service, then giving the post office a monopoly on first-class mail effectively prevents such cost information from being created and disseminated.

Each party needs to get its story out, at least in proportion to its standing. Each citizen needs access to news. Recent findings point up the problem: Asked how they are doing and expect to do, most Americans answer positively; asked how their country is doing, however, they say "terribly." My analysis is that each of us is the only possible expert on how we are doing but, by contrast, no one sees "the state of the country"; an opinion about the country is a synthetic, composite judgment that individuals are most likely to get from the major media (television networks, *New York Times, Wall Street Journal, Time, Newsweek*), because they are the only source of information that reaches most people. The media generally agree (except for the editorial pages of the *Wall Street Journal*) that the country is in big trouble. Without dwelling on media bias,[8] we see that communication corrupted by blatant distortion gravely impairs selection (an example might be the urgency in the media of removing asbestos from schools, which on net is harmful, not helpful, to health).

Four Conditions of Free Competition

Individual choice is also limited by the available parties and candidates, hence my previous stress on allowing freer entry to third, fourth, or fifth parties. It is neither easy nor necessarily desirable to mitigate the limitation on choice that stems from party organization of government. The voter may not be able to choose a list of favored policies because the vote is really between rival teams who come prepared with a list of preferences. In other words, voting is not like eating at a Chinese restaurant where you can choose courses from both group A and group B, but more like a French prix fixe menu in which the choice is only among complete dinners. Without cohesive parties, the voter may be able to act more perfectly expressive at the cost of losing the ability to give policy an across-the-board push in a desired direction.

[8]My opinion is that the major media share an eqalitarian bias (Wildavsky 1987, 1988a). The problem is not so much that there is bias (though I could live with a lot more CNN reporting), but that the bias is the same.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Principles of Size

Perhaps we can get a clue from the criteria that economists lay down for achieving perfect competition. Lists vary. George Stigler (1968, p. 181) says that the following four are most common. Taking them one at a time, we find first a size principle: "A main requirement of perfect competition is that the largest firm in a given industry is to make a trifling fraction of the industry's sales (or purchases)." The idea is to prevent a single firm from monopolizing a product or service so it could set prices without competition. In recent decades, however, thought and observation have revealed that it is much more difficult to establish an effective monopoly than had previously been thought, especially with the march of technology. The pressure and effects that facsimile machines exert on first-class mail, that fiber optics have on telephone and computer services, and that long-distance transmission and pipelines have on electric power are but a few contemporary reminders. Opinion is not yet unanimous but agreement is growing that only government can guarantee a monopoly. Thus, in principle, there is no obstacle to a competitive federalism even if one competitor is much larger than the others in certain fields.

But the federal government can run deficits, thereby supporting activities beyond the willingness of citizens to pay for them. This handicap (from the standpoint of a competitive federalism) may be overcome by customary or constitutional limits on revenue and/or expenditure. An equivalent safeguard may be obtained by offset or pay-as-you-go procedures. In general, such procedures, which now exist in the U.S. Senate, provide that (above a budget resolution that sets a spending ceiling), any proposed additional expenditure covering any of the 13 functions of the federal government must be accompanied either by cuts elsewhere or by new revenues. In this way, the doctrine of "opportunity costs" (defined as the value of an object being determined by what one has to give up to get it) is made manifest. Because federal budgeting would be shifted from all pleasure (indulging this or that group) to pleasure mixed with pain (some other worthy group will suffer), there is hope that the federal government will restrain its ambitions (see Wildavsky 1988c for vivid illustrations).

Another well-known strategy for imposing preferences without paying costs is to mandate performance by states or industry. The states have been trying to resist; after all, all senators and representatives come from states. But the road is uphill, partly because some states use the federal government to get others to share costs, thereby

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

avoiding the comparative disadvantage of paying while others do not; partly because it is hard for Congress to resist taking credit without paying; partly because of judge-made law; and partly for ideological reasons, to which I will return. It suffices to say here that if some judges were not persuaded that social justice requires greater equality of condition, the judges would not be usurping the legislative budget function.

In response, those who wish a more variegated federal system are trying to revive an old interpretation of the Ninth Amendment as a strict limitation of the federal government to only those matters specifically entrusted to it, all others being reserved to the states. I am dubious. For one thing, such a maneuver is bound to make the Constitution even more of a malleable political object, indirectly justifying expansionist efforts along opposite lines. For another, the so-far successful efforts to read egalitarian intent into the Constitution, by reading new rights into it unknown to its authors and ratifiers (Wildavsky 1990), suggests that the inevitable silences that accompany any governing document will be filled in by political forces. Inevitably, a constitution, if we are to live by it, must be reinforced or reinterpreted by the pulling and hauling of political forces. In other words, the competition among governments that we hope to establish must be a product of political competition, including political education, in the first place. In the end, it is the perceived morality of competition vis-à-vis the alternative of equalization that, as I shall argue in conclusion, will determine the result.

Principle of Independent Action

The second condition of free competition historically laid down by economists, according to Stigler (1968, p. 181), is that "firms are assumed to act independently" in competitive markets. Insofar as governments are concerned, the proviso must be interpreted to allow coalitions but to disallow coercion. States can get together with each other or the feds, but no one can order the others so that the recipient must obey. Certainly this condition is true among states. It is partly but not entirely true of the relations between the federal and state governments. The feds can make it next to impossible for the states to refuse its lead, but this is the exception. The rule is compromise, even rejection of federal rule. Should there arise strong national sentiment for uniformity, say in regard to environmental matters or welfare programs, however, it is doubtful whether this pressure could be resisted. I will presently discuss the grounds for uniformity. It is sufficient to reiterate that central governmental compulsion is irreconcilable with competition.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Principle of Knowledge

The third condition of perfect competition is that actors have "complete knowledge of offers to buy and sell." Partial to the Austrian subjective school of microeconomics myself, I think this criterion is right but overstated. After all, a basic aspect of entrepreneurship is the discovery of heretofore missing information of a kind that might not even have been available without the action of an enterprising economic agent. Study of command economies teaches me, however, the importance of market prices. Without them—that is, without widespread information on opportunities for buying and selling—much of the economy is limited to much less-efficient barter, as Adam Smith showed long ago.

Proponents of competitive federalism want to help make information available about the relative prices of governmental programs. Evaluations of programs should be multiple, varied, independent, and external. Multiplicity and variety are essential to mitigate bias. It is the residue of repeated evaluations, such as of farm subsidies or rent control, that are persuasive. Much more needs to be done, in my view, on such subjects as the cost of collecting taxes, the (dis)incentive effects, and generally what we are getting for what we are giving up. There is no magic in evaluations, which can be as contentious as the ideologies behind them. The trick is to present the costs and benefits in as revealing a manner as possible. Just as purchasing power parities are more revealing of standards of living than income per capita, so competition would be favored by creative comparisons of services versus taxes. The best comparisons are not cerebral but experiential: the ability to live in a state, to move, to try again, and so forth.

While another criterion, the divisibility of commodities, is not quite applicable to states, it is possible to distinguish states from each other. The greater the range of states' differences and the more distinctive each state, the easier it will be for citizens to choose among them. For large differences to be maintained, however, one must also accept inequality.

Principle of Movable Resources

The last condition-cum-criterion of perfect competition used by economists is that "resources are movable among users." Movability is more important than it might seem. If industry is required to use coal scrubbers so high-sulfur coal can be used, thus reducing use of low-sulfur coal, competition is ruled out. If Alaskan oil cannot be shipped to Japan, where it has the highest value, it must be sent where it can spill on American shores. For purposes of a competitive

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

federalism, impediments to the movement of people and commodities should be diminished. Restrictions on international trade are harmful in two ways: They weaken the selective mechanism, and they set a bad example for the federal system.

A Few Favorable Considerations, Maybe

A review of conditions and considerations relevant to competitive federalism suggests that while not ideal, they are sufficient. There are obstacles but not fatal impediments. And I think current events are uniquely favorable. The collapse of communism (with its overcentralized, command economy) in Eastern Europe and perhaps its transformation in the Soviet Union make federal features more relevant than ever. Whether the Soviet Union retains most of its components depends on its ability not to decentralize decisionmaking, which the USSR has already done to various so-called republics, but to find a formula for noncentralization.

It may appear that in 1992 or earlier the European Common Market will be moving in the opposite direction—toward centralization. And in certain respects that is true. But by no means all. On the contrary, a common currency is a prerequisite for enhanced competition. "A striking feature of contemporary Europe," Giandomenico Majone (1990) begins his prescient essay, "is the parallel movement toward greater political and economic integration, on the one hand, and toward regional decentralization on the other." Far from being at odds, "cultural diversity is the very essence of the European identity." As the Common Market exerts its homogenizing force, I agree with Majone, pre-existing regional diversity likely will assert itself.

What appears to be an inexorable movement toward a world economy paradoxically places a premium on flexibility. The ability to manufacture and deliver at short notice and to provide small batches of made-to-order goods is now advantageous. Thus, as Max Singer (1989) proposes for Japan (some people's ideal of an integrated national political economy), the center should be small with most matters handled at regional and local levels. In sum, one world implies many regions, and an even larger number of localities. And what better mode of self-directed coordination than noncentralized, competitive federalism.

But Federalism Is Incompatible with Egalitarianism

The one powerful countervailing trend is egalitarianism, as in greater equality of condition (see Wildavsky 1985). Like what? Like the movement to equalize the differences in power between men

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and women, black and white, gay and straight, animals and people, parents and children, rich and poor, and more.

Of course, egalitarians say they desire diversity. And, in a way, they do. But not diversity that gives some people advantages over others. By the time one has finished narrowing differences in result, however, there is not much room for noncentralization. Greater uniformity of result can be achieved only by national laws and regulations. Allowing, indeed encouraging, localities to differentiate themselves, indeed appealing on the basis of differences, is anathema to egalitarians. They love choice but not choice that will leave those they consider disadvantaged worse off. Since every difference is a form of inequality, the only ultimately tolerable result is no difference.

The development of affirmative action programs demonstrates the point. Initially, the idea was to beat the bushes for racial minorities and women. Then there were voluntary programs, followed by all but compulsory timetables and quotas. Initially, in universities, affirmative action was to be achieved by open enrollment. Who could say that entrance tests were perfect or that early grades precisely predicted later ones? Agreed. Let us admit all comers. But then higher dropout rates must be expected. No! Intolerable! The educational system is to blame. Perhaps if undergraduates saw more graduate students of their race or gender, they would do better. No, the whole campus is hostile! Let us downgrade Western civilization ("dead white men"). Let us make faculty and administration the same racial and gender proportion as the entire population.

The question has to be faced. A competitive federalism is based on the value of equal opportunity. Nationalization is based on uniformity. This apparent contradiction was resolved by the Jeffersonians and Jacksonians. They believed that opportunity and equality were compatible provided only that the federal government did not create artificial inequalities, such as grants of corporate charters, national banks, and national debt. Nowadays their successors believe the opposite: Only if the federal government redistributes resources so as to increase equality can there be true equal opportunity. No one should imagine that the passion for equality will stop short of the federal system.

The 21st century will usher in an intensified cultural-cum-ideological struggle over federalism. Will states become administrative units of a nationalized polity to enforce equality, or will they become competitors for services and allegiances in a competitive political economy? Will the peoples of the world, having observed that central command is a self-destructive way in which to run an economy into

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the ground, decide to apply the same winning formula to the polity? The "double security" referred to in this paper's title refers not only to federal and state relations but to the self-regulatory principles out of which the competing elements are constituted and renewed. Without an underlying belief in competition, including unequal results, format federal arrangements will become a mask for national uniformity.

References

Anton, Thomas. *American Federalism and Public Policy: How the System Works*. New York: Random House, 1989.

Buchanan, James M., and Brennan, Geoffrey. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press, 1980.

Clark, John, and Wildavsky, Aaron. *The Moral Collapse of Communism: Poland as a Cautionary Tale*. Unpublished manuscript, 1990.

Dake, Karl, and Wildavsky, Aaron. "Comparing Rival Theories of Risk Perception." *Daedalus* (forthcoming 1990).

Douglas, Mary. *Natural Symbols; Explorations in Cosmology*. London: Barrie and Rockliff, 1970.

Douglas, Mary, and Wildavsky, Aaron. *Risk and Culture*. Berkeley: University of California Press, 1982.

Elazar, Daniel. *The American Partnership: Intergovernmental Co-operation in the Nineteenth Century United States*. Chicago: University of Chicago Press, 1962.

Ellis, Richard, and Wildavsky, Aaron. *Dilemmas of Early American Presidents from Washington through Lincoln*. New Brunswick, N.J.: Transaction Press, 1989.

Ellis, Richard, and Wildavsky, Aaron. "A Cultural Analysis of the Role of Abolitionists in the Coming of the Civil War." *Comparative Studies in Society and History* 31 (1990): 89–116.

Friedman, Lawrence J. *Gregarious Saints: Self and Community in American Abolitionism, 1830–1870*. Cambridge: Cambridge University Press, 1982.

Fukuyama, Francis. "The End of History?" *The National Interest*, no. 16 (Summer 1989): 3–18.

Grodzins, Morton. *The American System*. New York: Rand McNally, 1966.

Hartz, Louis. *The Liberal Tradition in America: An Interpretation of American Political thought Since the Revolution*. New York: Harcourt, Brace, 1955.

Jennings, Kent, and Miller, Warren. *Parties in Transition*. New York: Russell Sage Foundation, 1986.

Lichter, S. Robert; Rothman, Stanley; and Lichter, Linda S. *The Media Elite*. Bethesda, Md.: Adler and Adler, 1986.

Majone, Giandomenico. "Preservation of Cultural Diversity in a Federal System." In *Comparative Constitutional Federalism*. Edited by Mark Tushnet. Greenwood Press (forthcoming 1990).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Malecha, Gary Lee. "Understanding Agrarian Fundamentalism: A Cultural Interpretation of American Populism." Paper prepared for 1988 Annual Meeting of the American Political Science Association, Washington, D.C., 1–4 September 1988.

Niskanen, William A. *Bureaucracy and Representative Government*. Chicago: Aldine, 1971.

Niskanen, William A. "Economies of Scale in the Provision of Public Schooling." Paper presented at the Sixth Annual Critical Issues Symposium, "Competition among Governments, Efficiency, and Economic Growth," Policy Sciences Program, Florida State University, 8–10 March 1990.

Ostrom, Vincent. "Polycentricity: The Structural Basis for the Emergence of Self-Governing Systems of Order." In *American Federalism: The Constitution of Order for a Self-Governing Society*. San Francisco: ICS Press (forthcoming 1990).

Polsby, Nelson, and Wildavsky, Aaron. *Presidential Elections*. 7th ed. New York: Free Press, 1988.

Popper, Karl. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Basic Books, 1962.

Singer, Max. "A Twentieth-Century Form of Government for Japan." Unpublished paper prepared for the Japan-U.S. Friendship Commission, 1989.

Sproule-Jones, Mark. "Urban Bads and the Structure of Institutional Arrangements." Unpublished manuscript, University of Victoria, Vancouver, B.C., 1979.

Stigler, George S. "Competition." *The International Encyclopedia of the Social Sciences*, Vol. 3. New York: Macmillan, 1968.

Thompson, Michael; Ellis, Richard; and Wildavsky, Aaron. *Cultural Theory*. Boulder, Colorado: Westview Press, 1990.

Webber, Carolyn, and Wildavsky, Aaron. *A History of Taxation and Expenditure in the Western World*. New York: Simon and Schuster, 1986.

White, Joseph, and Wildavsky, Aaron. *The Deficit and the Public Interest*. Berkeley: University of California Press, 1990.

Wildavsky, Aaron. "A Bias Toward Federalism: Confronting the Conventional Wisdom on the Delivery of Governmental Services." *Publius* 6 (July 1976): 96–120.

Wildavsky, Aaron. "Federalism Means Inequality." *Society* 22 (January/February 1985): 42–49.

Wildavsky, Aaron. "The Media's 'American Egalitarians.'" *The Public Interest* 88 (Summer 1987): 94–104.

Wildavsky, Aaron. "Where Bias and Influence Meet." *The Public Interest* 91 (Spring 1988a): 94–98.

Wildavsky, Aaron. "Resolved, That Individualism and Egalitarianism Be Made Compatible in America: Political Cultural Roots of Exceptionalism." Paper prepared for a conference on "American Exceptionalism" at Nuffield College, Oxford, 14–16 April 1988b.

Wildavsky, Aaron. *The New Politics of the Budgetary Process*. Glenview, Ill.: Scott, Foresman; Boston: Little Brown, 1988c.

Wildavsky, Aaron. "Robert Bork and the Crime of Inequality." *The Public Interest* 98 (Winter 1990): 98–117.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Policy Consequences of Intergovernmental Competition

Thomas R. Dye

Competition and Public Policy

Most debates over federalism are only lightly camouflaged debates over policy. Philosophers and economists (Rawls 1971, Buchanan and Tullock 1962) may assert a distinction between *constitutional* issues—deciding how issues should be decided—and *policy* issues—deciding the issues themselves. But in politics, constitutional decisions are never separated from policy outcomes. People *do* know what the policy consequences of various constitutional arrangements will be. Citizens as well as political leaders consistently subordinate constitutional questions to immediate policy concerns. Indeed, history is replete with examples of the same political leaders arguing one notion of federalism at one point in time to achieve their immediate policy goal, and then turning around and supporting a contradictory notion of federalism at a latter time when it fits a new policy goal. No American politician, from Thomas Jefferson onward, has ever so strongly supported a view of federalism that he ended up conceding a policy battle.

Debates about federalism must acknowledge policy consequences. Policies have been usefully classified as allocational, developmental, or redistributional (Lowi 1964). *Allocational* policies produce and distribute public goods and services to consumer-taxpayers. These policies encompass the provision of a broad range of state and local government services—education, health, welfare, streets and highways, police and fire protection, sewers, water and utilities, garbage disposal, parks and recreation, and so on. *Developmental* policies are those that directly enhance the economic well-

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is McKenzie Professor of Government at Florida State University.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

being of the state or community. These policies are directed toward economic growth; they include attracting industry, building transportation facilities, providing utilities, renewing urban areas, training the labor force for work, and so on. *Redistributional* policies are designed to redirect wealth to benefit particular segments of society to satisfy equity concerns. These policies include traditional welfare services, health care for the poor, unemployment compensation, and low-income housing, as well as progressive taxation. Note that these are analytical distinctions among types of policies; any specific government activity may have allocational, developmental, and redistributive elements within it.

Allocational Policy

Intergovernmental competition directly strengthens the allocative functions of government. Decentralization permits governments to match services with variations in demand. Greater overall citizen satisfaction can be achieved with multiple governments offering different packages of public services at different prices. Competition forces governments to become more efficient in their allocative activities, providing better services at lower costs. Competition forces government to be more responsive to citizens' preferences than monopoly government.

The major thrust of the original Tiebout model (Tiebout 1956) centered on allocational policies. And the value of intergovernmental competition in allocational policies has never been seriously challenged. Paul E. Peterson, one of the few political scientists to incorporate the notion of competition into a coherent theory of federalism, appears to agree:

> Allocation is the functional that local governments can perform more effectively than central governments, because decentralization allows for a closer match between the supply of public services and their variable demand. Citizens migrate to those communities where the allocation best matches their demand curve [Peterson 1981, p. 77].

Developmental Policy

Intergovernmental competition inspires state and local governments to be concerned with the impact of their taxing and spending policies on economic growth and to become directly involved in economic development activities. In recent years, international competition began to inspire similar concern in Washington. Prior to 1981, discussion of federal tax policy almost always centered on redistributional issues. Scholars and politicians who raised questions

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

about the growth-retarding effects of federal tax policies were ignored or ridiculed. But developmental issues have long been central to debates over taxing and spending in state capitols and city halls, and it is competition that has kept developmental concern in focus.

Monopoly governments have no direct way of estimating the growth-retarding effects of their tax policies. Cross-national comparisons are confounded by vast cultural differences, and longitudinal examinations of the effects of past tax policies are obscured by historical circumstances. Competing governments are in a better position to observe the economic consequences of their policy decisions.

Of course, we must distinguish between economic growth policy and state industrial policies. Economic growth policy is the provision of truly public goods and services with minimum burdens in taxes and regulations. State industrial policy generally involves special treatments—loans and subsidies, guarantees, and tax or regulatory exemptions—that in the long run contribute to economic inefficiency. But the effects of various industrial policy decisions are not easy to identify. The literature to date has failed to observe any significant aggregate growth effects for any direct state industrial development activities (ACIR 1967; Peirce, Hagstrom, and Steinbach 1979; Rasmussen, Bendick, and Ledebor 1984; Rubin and Zom 1985; Ambrosius 1986).

There are good theoretical reasons for predicting that direct state involvement in capital formation, as well as special tax treatments, will inevitably produce inefficiencies. But competition among the states in industrial policies is self-correcting in the long run. States will gradually learn from their mistakes as they compare their progress with other states. Consider how much worse a national industrial policy would be—a monopoly government allocating capital, dispensing subsidies, and granting special privileges and protections. (Not that the federal government does not already do so in innumerable tax code provisions and subsidy programs, but a national industrial policy would legitimize and enlarge the scope of these subsidies, privileges, and protections.) The only corrective to a national industrial policy would be international competition, and far more damage would be done to the nation before global competition would exert its correcting influence. Perhaps the results of the industrial policy experiments currently being conducted in America's "laboratories of democracy" will succeed in discouraging the federal government from pursuing national industrial policy.

Thus, development policies are well served by federalism. Peterson et al. (1986, p. 230) argue convincingly that the federal

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

government would be wise to leave economic development policy to state and local government:

> Since state and local governments are well equipped to pursue developmental objectives, most public efforts of this type should be left to them. By delegating responsibility for most developmental programs to state and local governments, the federal government would frankly admit its incapacity to use those programs to help populations with special needs.

Even strong proponents of state industrial policies warn against federal intervention. According to Osborne (1988, pp. 283-84):

> The America economy—indeed the world economy—is made up of a series of regional economies, each of which radiates out from a city or network of cities. Each regional economy is unique. Each has a different mix of industries a different labor market, a different set of educational institutions, and different capital markets. In this country, the governmental unit that most closely matches the regional economy is the state.

Redistribution Policy

The most serious challenge to intergovernmental competition arises in redistributional policy. Can multiple, competing governments undertake redistributive policies without creating unbearable free-rider problems for themselves? Will states and communities be restrained from providing the welfare services they would otherwise prefer because of the threat of an inundation of poor people from less-beneficent "free-riding" jurisdictions? Will each state and community wait for other states and communities to provide welfare services in the hope that their poor people will migrate to the more generous jurisdictions? Will tax burdens to support generous welfare services encourage the nonpoor—both households and business—to migrate to jurisdictions that impose lighter tax burdens because they provide frugal welfare services?

Welfare and Competition in Theory

Median Voter Model

Median voter models have provided economists and political scientists with simplified sets of assumptions about democratic political systems—assumptions that have expressly or implicitly formed the basis of a great deal of empirical research in state and local government. In the basic model, each individual consumer-taxpayer tries to maximize the tradeoff between various public goods and services and disposable private income. Each consumer-taxpayer balances

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the "utilities" obtained from higher levels of public services against the "disutilities" of higher taxes, in order to choose an "optimal" benefit level. After each consumer-taxpayer has made a choice based on his or her own preferences, tastes, needs, and so on, the choice of the median voter in any democratic political system should determine public policy. The "median voter" label is taken from the early work of Duncan Black (1958), who suggested that democratically elected governments would seek to make the median voter's indifference curve the state's indifference curve.

A simple version of the median voter model treats the political system itself as a neutral conversion mechanism, transforming the needs, demands, and preferences of the median voter into public policy. The model is similar to the early systems model in political science literature (Easton 1965, Dye 1966). The preferences of voter-taxpayers were usually specified in state policy studies in terms of median or average characteristics of the population—for example, median family income, median school level completed, or percentage of the population aged, young, black, urban, and so on. This required an inferential leap that characteristics of the populations determined median voter preferences and, hence, community preferences. When variations in public policies among the states were shown to correlate with variations in population characteristics, this simple median voter model gained empirical support. The traditional literature on the "determinants" of state and local government taxing and spending policies relied implicitly on this simple median voter model (Sachs and Harris 1964, Fisher 1964, Bahl 1965, Dye 1966, Hofferbert 1966).

Redistribution and the Median Voter

The median voter has an equity preference for redistribution. Few of us want to see poverty, hunger, homelessness, ill-health, or deprivation in our society, even if we do not expect to suffer these maladies ourselves. Moreover, some of us may wish to establish a floor for the human condition as an insurance policy against our own potential misfortunes. We may disagree on what minimal standards should be incorporated into this floor, but few would deny the existence of a general public interest in eliminating the worst conditions of poverty, or the greatest insecurities and afflictions of old age or disability, or the worse effects of discrimination.

The simple model of a median voter with a preference for some distributional government activities on behalf of poor, aged, young, and black populations can be developed further by assuming that these groups will contribute independently to the demand for redis-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tribution. Thus, increases in the poor populations should result in increased welfare spending through the direct political demands of the poor, as well as the charitable instincts of the median nonrecipient taxpayer. It turns out, however, that empirical tests reveal that increasing numbers and proportions of poor (and black) populations have the opposite effect. This suggests that the redistributional preferences of the median nonrecipient taxpayer vary inversely with the proportion of poor and black in the population (Dye 1966, Orr 1976).

Redistribution and Policy Responsiveness

The argument that intergovernmental competition disadvantages the poor rests on the assumption that competition itself forces state and local governments to reduce welfare benefits *below* those preferred by the median voter in their jurisdictions because of potential free-rider costs. States and communities are in economic competition with each other, and people, both rich and poor, can move freely from one jurisdiction to another. Governments are restrained from implementing the median voter's equity preferences for fear that so doing will attract poor people from other competing jurisdictions and repel affluent, productive, workers, firms, and investors.

Note that it is not sufficient that the proponents of this argument merely show that overall welfare benefits in the United States would be higher under a centralized government than they are under a competitive, federalism system. The true test of the effect of intergovernmental competition is whether it constrains the voter-preference-welfare-policy *relationship*, not whether welfare benefits would be raised by centralizing policymaking in Washington. Proponents of centralization must show that lower welfare benefits under decentralization reflect a *poorer match* between the true equity preferences of the median voter and public policy, than the match that would be obtained under centralized government.

My argument is that intergovernmental competition encourages policy responsiveness. Welfare policies, like all policies, are more responsive to citizen demands when undertaken by multiple competitive governments rather than monopoly government. *The effect of competition is neither to lower nor to raise welfare spending but to bring it into line with citizens' demands.* Competitive governments seek to match their welfare policies with both the compassion and the prudence of their citizens.

Competition and Redistribution

The argument that intergovernmental competition disadvantages the poor rests on three assumptions about human behavior—assumptions that are empirically verifiable albeit with some difficulty.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

First, the argument assumes that welfare benefits influence the decisions of the poor about where to live—that generous welfare benefits attract in-migrating poor and discourage resident poor from out-migrating in search of a better life.

Second, the argument assumes that affluent individuals, investors, and firms are repelled by the higher costs imposed by larger numbers of welfare recipients attracted by high benefits. By attracting "welfare seekers," states discourage productive labor and capital from coming to their states and may even encourage them to leave in search of lighter welfare burdens.

Third, the argument assumes that state and local policymakers act to avoid the free-rider problem—providing less redistribution than preferred by their median voter. This means that policymakers are aware of the benefit levels provided by sister states, and they use these levels as guidelines in their own decisionmaking. Fearing an influx of poor, they deliberately restrict welfare benefits to levels provided by their sister states, or perhaps even deliberately set benefit levels below those of their sister states to encourage the poor to leave their own state.

Interpreting the Evidence

Evidence relevant to these questions is important in discussions of American federalism—in determining the appropriate level of government for the provision of welfare services.

Early Research

The early research literature contradicted the assumption that the poor moved to maximize welfare payments. In a comprehensive review of this literature in 1974, sociologist Larry H. Long concluded that "no study has presented empirical evidence for the hypothesis that welfare payments themselves have attracted large numbers of persons to states and cities with high benefit levels" (Long 1979, p. 46). Long's own research reached the same conclusion. Poor people are less mobile than affluent people. When poor people move, they do so for family reasons or job opportunities. The migrating destinations of the poor are no different from the affluent: They tend to migrate to locations that promise employment opportunities. Knowledge of welfare benefit levels and eligibility rules in distant states or cities is very limited and is rarely a part of the decision of the poor to move.

Shapiro v. Thompson

The U.S. Supreme Court decision in *Shapiro v. Thompson* in 1969 altered the calculus of the poor in considering migration. The Court

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ruled that States could not constitutionally impose residency requirements on persons seeking welfare assistance. After 1969 welfare benefits became available to migrating poor in their new location as soon as their applications could be processed. Thus, it could be argued that earlier findings of the irrelevance of welfare benefits in the locational decision of the poor were no longer valid, that welfare benefits would have a more substantial effect on the migration of the poor after 1969.

Migration to Low-Benefit States

However, migration patterns in the United States over the past two decades confirm that people have been moving to perceived job opportunities and ignoring welfare benefits. Net migration has been away from the high welfare benefit states of the Northeast and Midwest and toward the low-benefit states of the South and West. Net migration in the United States is *inversely* related to welfare benefit levels in the states. The simple correlation between the proportion of new residents in the states and average monthly AFDC payments per family is $-.38$.[1] This relationship was reported in the early literature on labor migration (Galloway 1967).

It is possible, of course, that the poor could migrate opposite the net migratory flow. Net migration figures fail to tell us specifically about the migrating patterns of the poor. Several studies have attempted to disaggregate net migration figures, usually by race. These studies report that there are distinct differences between black and white migration patterns among the states. While the large white migration flow is toward states with *low* welfare benefits, black migration flows are toward states with *high* welfare benefits (Sommers and Suits 1973; Cebula, Kohn, Vedder 1973). These studies suggest that the poor do respond to welfare policy incentives by migrating to states offering higher welfare benefits. (For a summary, see Cebula 1979). These findings were offered some years ago as a justification for centralizing welfare policy (Orr 1976).

Higher Benefits and Immobility

A strong theoretical argument can be made that high benefits create immobilities among the poor. Generous benefits may weigh against the uncertainties of moving in search of job opportunities, thus causing a state to retain a large poverty population. While theoretically

[1]New residents are those who moved to the state in the previous five years, from 1980 Census of Population. Average monthly AFDC payments are these reported for 1980 by Social Security Administration's *Social Security Bulletin* and reprinted in U.S. Bureau of the Census's *Statistical Abstract of the United States 1982*.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

plausible, this thesis is difficult to study empirically. To learn why the poor do not move requires the causal study of a nonevent, always a difficult task.

Welfare Burdens and Migrating Wealth

Perhaps the weakest link in the argument that competition disadvantages the poor is the assumption that high welfare costs discourage economic growth. While theoretically plausible, it turns out that there is little convincing evidence that welfare costs raise tax levels to the point that affluent individuals, investors, or firms are encouraged to leave states or discouraged from coming to them. Welfare expenditures are only 13 percent of total state and local government expenditures. (Education is 35 percent of total state and local government spending.) And welfare spending burdens are *not* related to economic growth in the states. We might observe, for example, in Table 1 that the welfare burdens (total state-local welfare spending as a percentage of personal income) are not significantly related to economic growth (growth in per capita personal income), either over longer or shorter time periods. In contrast, highway (and infrastructure) spending is positively associated with growth in per capita personal income.

Of course, it is possible that the mere *perception* that welfare costs discourage economic growth may force policymakers in competitive states to reduce welfare benefits. We really need to see whether policymakers in the separate states eye each other in determining welfare benefit levels.

TABLE 1

STATE-LOCAL SPENDING BURDENS AND ECONOMIC GROWTH

Spending Burdens (Spending as a % of Income) at Beginning of Time Period	Growth in Per Capita Personal Income			
	1960–84	1960–70	1970–80	1980–84
Welfare	−.04	−.02	.12	.21
Education	.31	.11	.38*	.06
Highways	.31*	.07	.54*	−.13

NOTE: Figures are simple correlation coefficients for 50 states. An asterisk indicates a statistically significant relationship.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Variation in Benefits

Large variations among the states in welfare benefit levels suggests that decisionmakers do not give great weight to the benefit levels paid by other states in their own decisionmaking. If state comparisons had any significant effects on decisionmakers, we would expect to find greater convergence over time in levels of welfare benefits. No state would wish to "get out of line" in welfare benefits, fearing an influx of poor to generous states. Competing states would gradually move toward an equilibrium in benefit levels. But, in fact, differences among the states in welfare benefits are quite large and there is no trend toward convergence.

Average monthly AFDC payments in 1986, the principle state-administered welfare program, varied from highs of $532 in California and $507 in Minnesota to lows of $116 in Mississippi and $114 in Alabama, a difference of nearly 400 percent. Moreover, these great differences have persisted over time. The coefficient of variation for average monthly AFDC payments among the states has remained fairly constant since 1940. (See Table 2; see also Peterson and Rom 1989, Table 1; Dye 1990, Table 2–4.) Continuing wide variation in benefit levels strongly suggests that welfare policy in each of the states is made independent of other states.

Peterson's Argument

The most influential report purporting to show that American federalism disadvantages the poor is Peterson and Rom's recent analysis of the relationship between welfare benefits and the size of poverty

TABLE 2

CONTINUING POLICY VARIATION AMONG THE STATES, 1952–84

	Per Capita State and Local Government Expenditures					
	Total	Education	Highways	Welfare	Health	Total Taxes
1984	.45	.32	.55	.40	.32	.41
1980	.41	.32	.44	.36	.28	.21
1975	.28	.40	.42	.39	.38	.21
1970	.24	.20	.39	.43	.32	.21
1960	.20	.23	.34	.38	.31	.21
1952	.26	.44	.37	.49	.37	.23

NOTE: Figures are coefficients of variation: standard deviation divided by the mean. N = 50, except for 1952, when N = 48.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

populations in the states (Peterson and Rom 1989). They present evidence that an increase in welfare benefits results in an increase in a state's poverty rate. Specifically, they found that a change in the combined maximum cash and food stamp guarantee was a significant, positive, independent determinant of a change in the percentage of a state's population living below the poverty line.

But does the fact that higher welfare benefits result in more poverty in a state really prove that the poor were attracted from other states? Certainly a strong case can be made that increased welfare benefits can increase the size of the *within-state* poverty populations. A convincing argument can be made that generous benefits and relaxed eligibility rules encourage the formation of female-headed households, discourage recipients from obtaining entry-level jobs and the skills that these jobs could teach, and thereby increase poverty levels (Murray 1984). Increasing welfare benefits could increase the poverty population without any migration at all. Peterson and Rom do not present evidence that welfare benefits attract poor people to generous states.

Interestingly, Peterson and Rom report wide variations in welfare benefits among the states, yet continue to assert that the behavior of sister states motivates decisionmakers in state capitals. For evidence, they offer the fact that welfare benefits decline with increases in a state's poverty population. They dismiss a very plausible explanation of this relationship—that state policymakers tend to view welfare spending as a fixed proportion of state spending and, therefore, decrease benefits when the number of potential recipients increases in order to hold total costs relatively constant—because the savings from reduced benefits does not equal the costs imposed by a larger number of recipients. Instead, Peterson and Rom offer the less plausible explanation that "politicians, interest groups, and the news media may ask the question where poor people are coming from and what is bringing them" (p. 723). Presumably these actors then decide to reduce benefits to discourage more in-migration of poor people.

Summary

In short, the evidence to support the critical assumptions of the argument that competitive federalism disadvantages the poor is weak. Early studies suggested that the poor migrate for job opportunities and family reasons with little knowledge of welfare rules and payments in various jurisdictions. More recent studies suggest that the poor migrate opposite net flows and toward high benefit states. But even if we concede this assumption, there is no convincing evidence to support the notion that migrating poor and their welfare

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

burdens repel affluent individuals or firms. Generous welfare benefits may have a modest effect on the poverty rate, but that effect is more plausibly attributable to the direct and indirect poverty-inducing effects of transfer payments. Finally, continuing wide disparities among the states in welfare benefit levels, notably AFDC cash benefits, indicate that the convergence we would expect if policymakers took their cues from other states has not occurred.

Equity Preferences and Welfare Benefits

It is certainly true that the principles of efficiency and equity are frequently in conflict in political and economic affairs (Okun 1975). American federalism may involve some tradeoffs between these values. The political and economic benefits of intergovernmental competition may come at the costs of some equity considerations. If this is true, it should not necessarily doom the idea; society makes many such tradeoffs. Indeed, it is frequently argued that our society has traded off efficiency in so many of its redistributional programs that it threatens to impoverish both rich and poor. But let us examine the equity effects of federalism more closely. Perhaps it does not require a tradeoff of equity after all. Perhaps the poor will benefit from intergovernmental competition.

If the welfare policies of the national or state governments depended exclusively on the voting power of the poor, there would be very few welfare services in the United States. Welfare policies in the United States are primarily an expression of the equity preferences of middle-class Americans, not a response to the demands of the poor themselves. Indeed, when we examine the determinants of the welfare policies of the states, we find that the states with the greatest numbers and percentages of poor people provide the least in the way of welfare services. The states with the most generous welfare programs, in both absolute terms and in proportion to their incomes, are those with the larger personal incomes. The most important determinant of public welfare benefit levels is the income level of the population. Our equity preferences rise with increases in personal income. It is better to be poor in a rich state than in a poor one.

If multiple competing governments are seeking to maximize the true preferences of consumer-taxpayers, they must do so in the provision of welfare services as well as any other public goods. Is it reasonable to assume that migrating families and firms would wish to locate in areas of squalor, poverty, homelessness, and human degradation simply to avoid the costs of welfare services? The cross-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sectional evidence—that wealthy states provide more welfare services than poorer states—suggests that welfare services are preferred by many middle-class residents and wealthy firms.

The view that competitive federalism would result in the underproduction of welfare services rests largely on the belief that the poor will migrate to the most generous jurisdictions. But the poor are the least mobile segments of society. Whatever their impact on state welfare services, it is certainly less than the impact of mobile upper-middle-class Americans and wealthy firms. Under competitive federalism, the most mobile sectors of the population and the economy exercise the greatest influence over public policy. And there is convincing evidence that these affluent segments of the population are more likely to support welfare services than society generally.

In short, intergovernmental competition tends to ensure policy responsiveness, Admittedly, it is not designed to maximize equity, that is, to encourage redistributional policies. But the preferences of consumer-taxpayers are not necessarily narrow and selfish. Competitive federalism ensures responsiveness to redistributional preferences, as well as other policy preferences.

Matching public policy to citizen preferences is the essence of responsive government. Competitive elections and political parties were designed to achieve this goal. But decentralized government is also a way to match citizen preferences with public policies. Given variation in policy preferences across the nation and given independent governments with the authority to offer a wide range of policies at different costs, a truly federal government can achieve a far better match between preferences and policies than a centralized government imposing uniform policies throughout the nation.

References

Advisory Commission on Intergovernmental Relations (ACIR). *State-Local Taxation and Industrial Location.* Washington, D.C.: ACIR, 1967.

Ambrosius, Margery M. "Effects of State Economic Development Policies on the Health of State Economies." Paper presented at the Midwest Political Science Association, Chicago, 1986.

Bahl, Roy W. "Determinants of Changes in State and Local Government Expenditures." *National Tax Journal* 18 (March 1965): 50–57

Black, Duncan. *Theory of Committees and Elections.* Cambridge: Cambridge University Press, 1958.

Buchanan, James M., and Tullock, Gordon. *The Calculus of Consent.* Ann Arbor: University of Michigan Press, 1962.

Cebula, Richard J. "A Survey of the Literature on the Migration Impact of State and Local Government Policies." *Public Choice* 1 (1979): 69–83.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Cebula, Richard J.; Kohn, M.; and Vedder, Richard K. "Some Determinants of Interstate Migration of Blacks." *Western Economic Journal* 11 (December 1973); 500–5.

Downs, Anthony. *An Economy Theory of Democracy*. New York: Harper and Row, 1957.

Dye, Thomas R. *Politics, Economics, and the Public*. Chicago: Rand McNally, 1966.

Dye, Thomas R. *American Federalism: Competition among Governments*. Lexington, Mass.: Lexington Books, 1990.

Easton, David. *A Framework for Political Analysis*. Englewood Cliffs, N.J.: Prentice Hall, 1965.

Fisher, Glenn W. "Interstate Variation in State and Local Government Expenditures." *National Tax Journal* 17 (March 1964): 57–74.

Galloway, Lowell E. "The Economics of Labor Mobility." *Western Economic Journal* 5 (June 1967): 211–23.

Hofferbert, Richard. "The Relation Between Public Policy and Some Structural and Environmental Variables in the American States." *American Political Science Review* 60 (March 1966): 73–82.

Long, Larry H. "Poverty States and Receipt of Welfare among Migrants and Nonmigrants in Large Cities." *American Sociological Review* 30 (February 1979): 46–56.

Lowi, Theodore J. "American Business, Public Policy, and Political Theory." *World Politics* 16 (July 1964): 677–715.

Murray, Charles. *Losing Ground: American Social Policy, 1950–1980*. New York: Basic Books, 1984.

Okun, Arthur M. *Equality and Efficiency: The Big Tradeoff*. Washington, D.C.: Brookings Institution, 1975.

Orr, Larry L. "Income Transfers as a Public Good." *American Economic Review* 66 (June 1976): 359–71.

Osborne, David. *Laboratories of Democracy*. Boston: Harvard Business School, 1988

Peirce, Neal R.; Hagstrom, Jerry; and Steinbach, Carol. *Economic Development: The Challenge of the 1980's*. Washington, D.C.: Council of State Planning Agencies, 1979.

Peterson, Paul E. *City Limits*. Chicago: University of Chicago Press, 1981.

Peterson, Paul E.; Rabe, Barry G.; and Wong, Kenneth K. *When Federalism Works*. Washington, D.C.: Brookings Institution, 1986.

Peterson, Paul E., and Rom, Mark. "American Federalism, Welfare Policy, and Residential Choices." *American Political Science Review* 83 (September 1989): 711–28.

Rasmussen, David; Bendick, Bart; and Ledebor, Larry. "A Methodology for Selecting Economic Development Incentives." *Growth and Change* 15 (January 1984): 18–25.

Rawls, John. *A Theory of Justice*. Cambridge: Harvard University Press, 1971.

Rubin, Barry M., and Zom, C. Kurt. "Sensible State and Local Economic Development." *Public Administration Review* 45 (March/April 1985): 333–39.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Sachs, Seymour, and Harris, Robert. "The Determinants of State and Local Government Expenditures and Intergovernmental Flow of Funds." *National Tax Journal* 17 (March 1964): 75–85.

Sommers, Paul M., and Suits, Daniel B. "Analysis of Net Interstate Migration." *Southern Economic Journal* 40 (October 1973): 193–201.

Tiebout, Charles M. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (October 1956): 416–24.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Interstate Tax Competition, Incentives to Collude, and Federal Influences

Bruce L. Benson

Incentives to collude exist in every competitive market. Nonetheless, the successful formation of cartels among private-sector firms is rare, in part because the cartel must overcome the costs of organizing and then of policing the members' agreement. This difficulty to form a cartel has frequently led organized groups of firms to turn to an authority with coercive power (government) for support. Indeed, as George Stigler (1971) explained, economic regulation is most clearly understood as a process by which government established legal cartels by fixing prices, assigning quotas, and limiting entry. Even in the case of cartelization through regulation, however, rent dissipation occurs in the process of rent-seeking competition (e.g., Tullock 1967, Posner 1975, Tollison 1982, Benson 1984) and nonprice competition (e.g., Stigler 1968, Douglas and Miller 1974, Fournier 1985).

The purpose of the following discussion is to extend Stigler's economic theory of regulation, with its roots in the theory of the firm and public choice theory, by illustrating that similar kinds of incentives to collude exist for competing governments.[1] However, as with private firms, successful collusion is generally thwarted by high costs of

---

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at Florida State University. Parts of this paper draw from the author's work with Ronald N. Johnson (see Benson and Johnson 1986a, 1987).

[1]Vedder (1990) discusses interstate tax competition in the context of a theory of the firm/public choice model, and McKenzie and Staff (1981) propose an extension of Stigler's theory of economic regulation to include government units in the context of their discussion of revenue sharing. The presentation that follows extends Vedder's analysis by considering the incentives to collude, and the consequences of nontax-rate competition. It also involves a more extensive analysis of the form and consequences of interstate competition than appeared in McKenzie and Staff and consideration of efforts to avoid competition beyond just revenue sharing. In the process, certain oversights by McKenzie and Staff (1978, 1981), as well as by Friedman and Kurth (1981), will be rectified.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

organizing and policing. Consequently, any relatively long-term success typically requires the assistance of a higher authority with some coercive power over the government units attempting to form a cartel. These arguments are developed in the context of an examination of interstate competition for business investment and tax base.

Politics and State Tax Rates

The impact of state and local taxes on capital investment are examined in Benson and Johnson (1986b) where interstate tax competition is accounted for by using a relative tax-rate measure. Furthermore, the econometric procedure was designed to reflect the possibility of a lagged impact of changes in relative taxes on capital formation and to detail the time profile of the lag. Our results indicated a statistically significant lagged response. The major portion of the long-run effect of relative tax changes on capital investment occurs within three to four years of the change. The estimates indicate that the mean long-run elasticity of a change in relative taxes on relative capital expenditures is -1.02.[2]

The finding of a significant distributed-lag effect of state and local taxes on economic activity is important in the context of a theory of the firm/public choice explanation of interstate tax changes and differences. Dwight Lee and James Buchanan (1982) argue that political decisions are made on the basis of a limited time perspective. Politicians are seen as having little motivation to consider consequences much beyond their next re-election effort. These incentives are heightened by the notion that voters tend to be rationally ignorant about the long-run consequences of tax policies.

In the context of the findings in Benson and Johnson (1986b), a politician who increases tax rates to provide additional benefits to supporters, and who argues that the increases will not affect economic growth, may have a short-term re-election advantage over an opponent who contends that such tax increases will have a significant negative impact on economic activity. By the time the evidence is accumulated, the election may be over. As Lee and Buchanan (1982, p. 354) noted, "so long as government makes its fiscal decisions on the basis of a time horizon shorter than the period required for full taxpayer adjustment to tax rate changes, observed tax rates will be higher than those that a far-seeking or 'enlightened' government would impose."

Roughly half of the negative impact of an increase in relative taxes occurs within two years (Benson and Johnson 1986b), however; thus,

[2]Also see Vedder (1990) and references cited therein.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

state and local tax policy should be constrained to some degree by interstate tax competition. Interstate tax competition can be expected to moderate the tendency toward excessively high state and local taxes, and to narrow the distribution of tax levels across states. This competitive impact can be seen by considering what happens when a state gets too far out of line with competing states.

According to the Advisory Commission on Intergovernmental Relations (ACIR 1981), for example, the tax burdens of New York and Massachusetts got out of line by the mid-1970s, relative to their neighbors. New York's public sector outlay per dollar of personal income was about 30 percent higher than the rest of the nation by the middle of the decade, for instance. As a consequence, among the 303 major manufacturing firms within the food, textiles, apparel, fabricated materials, machinery, and electrical equipment industries that relocated between 1969 and 1976, 91 left New York (ACIR 1981, pp. 46–47). Of those, 29 moved to other states in the Mideast. Massachusetts lost 25 major manufacturing establishments to states within the New England region alone (17 moved to New Hampshire) during the period. The experience of New York and Massachusetts suggests that if a state government raises taxes without regard to tax levels in other states, that state will find itself growing at a substantially slower rate within a few years. As a consequence, states are likely to react by offering more competitive taxes. Both New York and Massachusetts did exactly that. In 1979, New York Governor Hugh L. Carey wrote: "During the past four years, New York State has demonstrated that the way to provide tax relief and revive the economy is to control the growth in spending. Spending increases have been held far below the rate of inflation, making tax cuts possible three years in a row" (ACIR 1981, p. 29). Indeed, the rate of increase in New York's state and local tax burden (1.1 percent) was below the national average (1.2 percent) for 1975–78, and, perhaps more importantly, it was considerably below Connecticut's (2.5 percent) and New Jersey's (2.3 percent) increases.

The concept of tax-rate competition among states does not imply that states will have identical overall tax rates. This is not a perfectly competitive market. It is more appropriately described as oligopolistic. But, to the extent that interstate tax competition interferes with the objectives of politicians, we can anticipate attempts by public officials to avoid competition (i.e., collude).

Efforts to Reduce Interstate Tax Competition

There are many examples of efforts by state governments to collude. The governors of six Great Lakes states, for instance, discussed

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the possibility of "calling a truce on attempts to pirate industry from each others' states" (Carlson 1983, p. 31). However, the six governors were unable to reach an agreement. Minnesota's governor took the lead; after unsuccessful attempts to convince governors from Iowa and the Dakotas to reduce competition for industries, he reached an agreement with Wisconsin. The governors of Minnesota and Wisconsin publicly pledged not to initiate any state-sponsored business recruiting in the other's state, although contacts by firms from the other state could still be pursued. Governors from Indiana and Illinois refused to join such a pact, however, despite interest by Michigan and Ohio.

Even such a limited agreement is not likely to be effective. Firms that consider relocating or branching could still initiate contact with state officials, and negotiations for tax concessions and financial inducements could follow.[3] Furthermore, the "no-raid" agreement did not put any limit on direct tax-rate competition, nor was it an agreement to set tax rates analogous to a price-fixing cartel. Yet even this extremely limited agreement could not be negotiated with most of the states in the region. This lack of cooperation is not surprising. Private sector cartels are extremely tenuous arrangements if they can be achieved at all. The same appears to hold true of public cartels. As Robert Kleine (1977, p. 170) noted, "At the state level, tax competition could be reduced by cooperative efforts of regional economic planning groups such as the Southern Growth Policy Board (SGPB). Neither the SGPB nor any of the other regional groups currently in existence has taken steps to reduce competition among member states."

Under certain circumstances it might be anticipated that collusion should be relatively easy. For instance, Montana and Wyoming have huge deposits of low-sulfur coal. It is by far the cheapest coal in the

[3]This point brings up an important issue that pertains to much, if not all, of the literature on state and local taxes, location choice, and economic growth. The literature includes Vedder (1990) and Benson and Johnson (1986a, 1986b), and treats the taxes as endogenous to the firms making location choices. Actually, firms may have a significant impact on the taxes they pay. They can bargain directly for tax breaks, and perhaps more importantly, they can lobby for changes in tax rates. Thus, the most attractive site in terms of tax considerations, *ceteris paribus*, may not be the location with the lowest current tax rates, but rather, it may be the location with the political system that appears to be the easiest to influence with regard to future taxes.

Tax rates may be negatively correlated with determinants of political ability to influence, of course, but the correlation is clearly far from perfect. New York is a high tax-rate state, for example, but it also has a very substantial tax concession program. Newly locating or highly mobile firms may be able to obtain very favorable tax treatment from the state despite the fact that current residents and relatively immobile firms pay high taxes.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

United States to produce, so even when transportation costs are added, the coal competes very favorably in the Midwest and Southwest markets of the United States (Kolstad and Wolak 1983, p. 449). Furthermore, excise taxes on such coal are largely exported, so citizens of the states find such taxes to be quite desirable, and the source of the tax—coal—is highly immobile so it cannot flee (although coal producers had the choice of developing the reserves in Montana or Wyoming, or of using more expensive coal elsewhere). When Montana raised its severance taxes of surface-mined coal to 30 percent, Wyoming could have followed suit, tacitly colluding, or it could have taken advantage of its competitive position in attracting coal development and left its taxes alone. Wyoming responded by raising its taxes, but by substantially less than Montana had. Charles Kolstad and Frank Wolak (1983, p. 455) contended that Montana and Wyoming could actually have charged substantially higher taxes than they did, through collusion; instead, there were "substantial losses in market power as a result of interregional competition, but despite these loses, . . . [the two states] captured sizable amounts of rents".[4] Even when the situation appears ripe for interstate tax collusion (i.e., a small number of competitive states, an immobile tax base, and an exportable tax), it does not appear to materialize.[5] Taxes may be higher than they might be in a more competitive environment, but they do not approach the perfectly collusive level.

Private sector firms that want to limit competition will typically turn to the government for help (Stigler 1971). Similarly, several organizations of state and/or local governments exist for the primary purpose of lobbying the federal government. These lobbying groups include the Northeast-Midwest Economic Coalition of more than 200 members of Congress from 18 states, the Western Governors Association with governors from 14 states and 3 territories, the National Governors Association, the U.S. Conference of Mayors, the National League of Cities, and so on.

Of course, there is little that the federal government can do, constitutionally, to reduce interstate tax competition directly. Congress and the courts rarely interfere with the taxing powers of sovereign states. Nonetheless, the federal government has established institutional arrangements that can significantly curtail the impact of inter-

[4]Zimmerman and Alt (1981) and Church (1981) also estimated that the cartel tax rate on such coal substantially exceeded the actual tax rate.

[5]OPEC successfully formed a cartel under similar circumstances, of course, but with the added political impetus of the Arab-Israeli conflict. However, this cartel was successful for only a relatively short time until internal cheating (exceeding of quotas) and entry into the market (North Sea oil, etc.) made the cartel ineffective.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

state tax competition and can reduce incentives to compete. Two factors of this type stand out. One is deductibility of state and local taxes from federal taxes.[6] The other is the growing share of state and local expenditures that are covered by federal revenues. Let us briefly consider each of these factors.

Federal Deductibility of State and Local Taxes

The allowance of state and local tax payments as business expenses for federal taxes, along with the deductibility of state and local income and property taxes from federal personal income taxes, effectively reduces real state and local tax rates to well below their nominal rates. In addition, such deductibility substantially narrows the actual tax differentials between states (ACIR 1981, Benson and Johnson 1986a, Vedder 1990).

Table 1 provides some evidence of this effect. In 1977, the range of total personal taxes relative to the average for a person with $50,000 in income, when state and local taxes alone were considered, was from .53 in Texas to 1.63 in New York. However, when federal taxes were added in, the range narrowed, with Texas at .90 and New York at 1.13.

The deduction of state and local taxes substantially diminishes the cross-state variation in taxes by transferring much of the federal tax burden from residents of high-tax states to residents of low-tax states. It also reduces the incentives to hold tax rates down. For instance, it seems reasonable to expect that New York and other high-tax states would be forced to make tremendous tax cuts to prevent substantial economic decline if deductibility were eliminated. On the other side of the coin, Texas's apparently competitive tax policy actually reduces personal taxes only by a relatively small percentage below the average, because Texas citizens have had fewer deductions from their federal taxes. In other words, federal deductibility substantially reduces the competitive impact of a low-tax policy and reduces the incentives to compete in terms of tax rates. For every dollar in revenue that state politicians give up, the state's citizens and businesses gain less than a dollar in overall tax relief.

It might be noted that federal deductibility of state and local taxes apparently was not solely designed to mute interstate tax competition. Its purpose supposedly was to guarantee that individuals' total tax rates would remain below 100 percent. In fact, however, interstate tax competition would guarantee this anyway, and the actual impact

[6]Also see Benson and Johnson (1986a) and Vedder (1990) for discussions of the impact of federal deductibility on interstate tax competition.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 1

COMPARISON OF 1977 PERSONAL TAXES IN SELECTED STATES, $50,000 INCOME

States	State and Local Taxes as Percent of Income	Index[a]	State, Local, and Federal Taxes as Percent of Income	Index[a]
Arizona	7.7	.74	27.9	.95
California	12.0	1.14	30.4	1.03
Colorado	8.7	.83	28.5	.96
Connecticut	8.5	.81	28.4	.96
Georgia	9.2	.88	28.7	.97
Illinois	8.0	.76	28.0	.95
Indiana	8.8	.84	28.5	.97
Kentucky	9.6	.91	29.0	.98
Maine	12.0	1.14	30.4	1.03
Massachusetts	14.4	1.37	31.8	1.08
Michigan	9.2	.88	28.8	.97
Minnesota	14.8	1.41	32.0	1.08
New Hampshire	7.0	.67	27.5	.93
New Jersey	13.3	1.26	31.1	1.05
New York	17.1	1.63	33.4	1.13
North Carolina	8.3	.79	28.2	.96
Ohio	8.1	.78	28.2	.95
Pennsylvania	11.8	1.12	30.2	1.02
Rhode Island	12.6	1.21	30.8	1.04
Texas	5.6	.53	26.7	.90
Vermont	10.0	.95	29.2	.99
Wisconsin	14.1	1.34	31.6	1.07
Average	10.5	1.00	29.5	1.00

[a]Total personal taxes relative to the average.
SOURCE: ACIR (1981).

of deductibility appears to be that total tax rates are higher than they would otherwise be. Furthermore, whether or not muting interstate competition was the purpose of deductibility, states certainly have taken advantage of the situation. High-tax states have been able to shift part of their tax burden onto the federal tax base. That is why proposals in the early 1980s to eliminate deductibility met with such intense lobbying in opposition, particularly from politicians in relatively high-tax states. Deductibility protects those states from competition, at least to a degree, and elimination of deductibility would force them to lower their tax rates substantially or face signifi-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

cant losses as business activity shifts to other states. Politicians from low-tax states also benefit from deductibility in the sense that it reduces the incentives to compete in terms of tax rates, thus possibly allowing them to set rates that are higher than they would otherwise be.

Fiscal Federalism

In 1961, 11.1 percent of total state and local revenues were transfers from the federal government. The portion of state and local revenues coming from the federal government rose throughout the 1960s and 1970s. By the 1965–66 fiscal year, this value had reached 13.4 percent, and the 1970–71, 1975–76, and 1979–80 figures were 15.7, 18.2, and 18.4 percent, respectively. (Reagan administration policies reversed this trend, as discussed below.) Total federal transfers to state and local governments grew from 1.4 percent of GNP in 1959 to 3.3 percent in 1980. This growing portion of revenues coming from the federal government should be considered in the context of growing state and local tax burdens and of increasing variance in taxes across states (see Vedder 1990 and Table 2). State and local government lobbying has played a major role in the development of

TABLE 2

MEAN AND STANDARD DEVIATION OF TAX RATES FOR 48 STATES, 1960–81

Year	Mean	Standard Deviation	Year	Mean	Standard Deviation
1960	.097	.014	1971	.117	.014
1961	.098	.012	1972	.122	.015
1962	.101	.012	1973	.123	.015
1963	.100	.011	1974	.118	.016
1964	.103	.011	1975	.116	.015
1965	.105	.012	1976	.120	.016
1966	.106	.012	1977	.121	.017
1967	.105	.012	1978	.121	.017
1968	.107	.013	1979	.117	.015
1969	.110	.013	1980	.112	.015
1970	.114	.014	1981	.111	.015

NOTE: The tax rate is defined as total state and local revenue divided by state personal income. Tax revenue data are from U.S. Department of Commerce, Bureau of the Census, *Governmental Finances*, various years. State personal income is from U.S. Department of Commerce, Bureau of Economic Analysis, revised data, August 1983.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the system wherein federal tax revenues are used to cover state and local expenditures.

There are "public interest" explanations for state government lobbying efforts to obtain federal transfers, of course. For example, it has been argued that economies of size exist in tax collection and administration. Thus, low-cost federal collection followed by federal transfers (e.g., revenue sharing, grants-in-aid) to state and local governments will supposedly allow those governments to lower taxes. However, this argument has been questioned. There is evidence indicating that the primary goal behind lobbying for federal transfers has been to increase state and local expenditures. For example, a 1977 study of revenue sharing, which actually has been a relatively small component of total federal transfers, concluded that more revenue-sharing dollars were used to maintain or expand programs than to reduce taxes (Juster 1977). Only 20 to 25 percent of revenue-sharing funds went toward tax abatement uses during the first few years of the program's existence (it was started in 1972). Similarly, a study of the impact of grants-in-aid concluded that only 15 percent of the grants were used for tax relief (McGuire 1979). The implication is that total taxes from federal, state, and local sources collected to cover state and local expenditures are substantially higher than they would have been without such transfers (as explained below, these results are relatively short-lived because incentives to compete still exist; the form of the competition simply changes).

Thus, just as a successful private-sector cartel, established and policed through regulation, raises revenues through a centralized pricing policy, federal centralized tax collection and subsequent distribution to the states tend to raise total tax revenues and state government expenditures. But the implications of the availability of large amounts of federal funds for state and local use go beyond the immediate impact that studies, such as those just cited, are able to detect. At the same time as the tax collecting function has increasingly been centralized and performed by the federal government, the variance-reducing impact of deductibility has increased. Therefore, incentives to lower state and local tax rates for competitive purposes have declined.[7]

On June 30, 1981, President Reagan told the National Conference of State Legislatures that the ultimate objective of his federalism policy initiatives was "to use block grants as a bridge leading to the day when you will have not only the responsibility for programs that properly belong at the state level, but you'll have the tax sources now

[7]See McKenzie and Staff (1978, 1981) for similar conclusions regarding revenue sharing.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

usurped by Washington returned to you—ending the round trip of the people's money to Washington and back minus carrying charges."[8]

Indeed, the share of state and local revenues coming from the federal government fell from its 1979–80 high of 18.4 percent to 14.7 percent in 1984–85 and 13.6 percent in 1986–87. Significantly, however, this percentage remains above what it was in the mid-1960s, *and* even though the percentage has fallen, the actual dollar value of federal transfers fell only during the first year of the Reagan administration (from $90.3 billion in 1980–81 to $86.9 billion in 1981–82); thereafter, federal transfers rose continuously throughout the Reagan period. Federal transfers to state and local governments reached nearly $115 billion in 1986–87, for a 32.3 percent increase over five years. Furthermore, the relative reductions achieved by the Reagan administration were part of a political deal wherein substantial reductions in federal controls on state uses of funds were made at the same time. With the reductions in federal controls, state and local officials had more discretion in the use of the transferred funds and were more willing to accept the relative decline in federal monies.

The Economic Regulation Analogy Revisited

The Lee-Buchanan view of the incentives that government officials have to set relatively high tax rates is an important consideration in the context of this discussion, because the theory of the firm analogy is imperfect. For instance, McKenzie and Staff (1978) contend that state governments are likely to reduce the quantity and quality of goods and services with a reduction in competition accompanying increased revenue sharing. This result is certainly expected in a profit-seeking, private-sector, regulation-created cartel. But elected government officials presumably do not restrict output to raise profits. They are more likely to increase revenues to *increase* the quantity and quality of benefits that they provide to their supporters, thereby enhancing their re-election opportunities.

Another point of confusion created by the theory of the firm analogy involves the McKenzie-Staff (1981, p. 373) prediction that tax variances among states should decrease when interstate competition is muted. In fact, variance has increased as interstate competition has been muted (see Table 2). Price variance may diminish under certain circumstances as a private sector cartel becomes effective (variance might also fall as competition is increased; however, parallel action can reflect either competition or collusion). Nonetheless, federal

[8]Quoted in the *National Journal*, 22 August 1981, p. 1492.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

muting of interstate tax-rate competition enables politicians in states with relatively high taxes to keep their taxes high and even raise taxes with relatively less fear of significant consequences, thus tending to increase variance. Simultaneously, there is a reduction in the competitive impact of low taxes and, therefore, in the incentives to hold taxes down in low-tax states, thus tending to reduce variance.

The net impact of these two anti-competitive forces on variance is not predictable. However, if, as Vedder (1990) hypothesizes, the elasticity of demand for government services is higher in traditional low-tax states than in traditional high-tax states, then the variance should increase. This is clearly the case when we compare the 1970s to the 1960s. Vedder's (1990) results indicate that changes in deductibility and in federal transfers that occurred during the early Reagan years may have reduced interstate tax variance, but that these changes were short-lived. This result may reflect the continued increase in the absolute level of federal support, but it also could reflect other institutional developments that typically accompany cartelization.

Nontax Competition

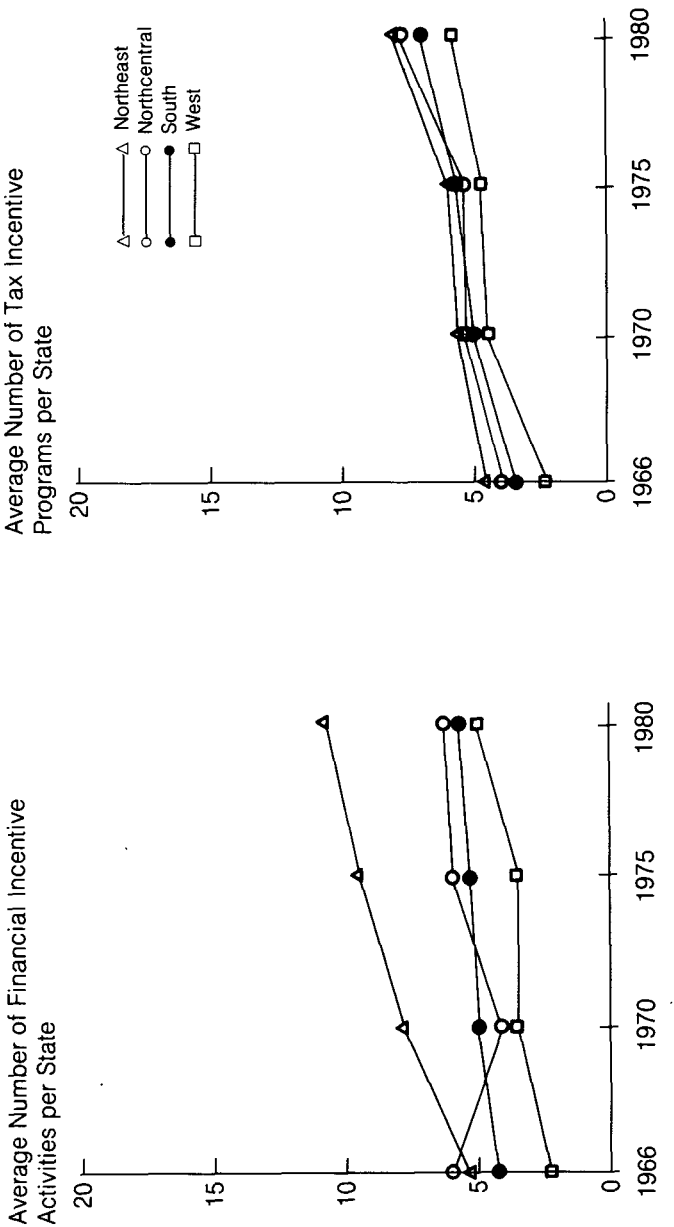
Regulatory-generated cartel rents in private markets tend to be competed away through nonprice competition (Stigler 1968, Douglas and Miller 1974, Fournier 1985). The same type of phenomenon has occurred in the intergovernmental competition to attract tax base. As federal deductibility and federal transfers have worked to mute the incentives to lower tax rates for competitive purposes, interstate competition has taken on a different form. Nontax-rate types of competition have become commonplace.[9] There has been a substantial increase in the use of tax and financial incentives by states to attract and hold business firms. Between 1966 and 1980, virtually every state introduced at least one new financial incentive, tax exemption, or special service in an effort to attract or retain industry (Miller 1983). During that period, 34 states added at least 10 new direct-assistance programs, and the number of such programs more than doubled in 28 states. Initially, these programs tended to be directed toward attracting new firms or branches, but recently they have been used to encourage expansion of firms already located in the state. Figure 1 details the upward trend in the availability of such programs by census region.

[9]There is another form of competition as well. States now compete for federal transfers much as other interest groups compete in the rent-seeking process.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FIGURE 1

GROWTH OF STATE FINANCIAL AND TAX INCENTIVE PROGRAMS BY CENSUS REGION



SOURCE: Miller (1983, p. 5).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The spread of such programs follows the same competitive process as tax-rate competition. State politicians prefer not to compete at all, of course, and federal institutions have reduced the pressure to compete on the basis of overall tax rates. Nonetheless, state and local officials have strong incentives to devise policies that can be selectively applied in cases where they may make a difference. But once one state introduces such a policy instrument, its neighbors meet the competitive pressure with an instrument of their own to offset the competitive disadvantage. The ultimate consequence is that states are unable to avoid competition, since each state keeps its tax concession and financial inducement program in line with its competitors. The ACIR (1981, p. 5) concluded that "widespread state enactment of incentives to businesses in the 1960s and 1970s has tended to neutralize their pulling power." Tax rates appear to vary more across states because, as a consequence of federal activities, these rates do not matter to politicians as much as they otherwise would. However, competition has simply taken on a different form, and the theory of the firm/public choice model still explains the process.

In this light, it should be clear that overall state and local revenues are not likely to rise above the "competitive" level in the long run, even when federal programs mute the effectiveness of one avenue of competition. Findings by Martin McGuire (1979) and Thomas Juster (1977) suggest that grants-in-aid and revenue sharing have increased state and local revenues. However, David Friedman and Michael Kurth's (1981) examination of revenue sharing suggests that this is a short-lived occurrence. In particular, they contend that total receipts (taxes plus revenue shared) exceeded what taxes would have been without revenue sharing between 1972 and 1976. But after 1976, tax collections were reduced by the full amount of the revenue-sharing funds (note that Juster's study was based on data up to 1977 so his results do not contradict Friedman and Kurth). Tax collections fell because of nontax-rate competition, just as rents in private sector cartels are dissipated. As Friedman and Kurth noted, "getting the monopoly solution is not as simple as fixing price" (1981, p. 365).

Note that the fact that revenue-increasing consequences of federal programs are short-lived does not imply that lobbying by state and local officials to establish such programs is irrational, even if they can accurately predict this result. Given the relatively short time horizon of elected public officials, an action is desirable if spending can be increased for even a short period without increasing taxes.[10]

[10]This conclusion also implies that state and local officials benefit most from new

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Federal muting of tax-rate competition does cause long-term changes, even though state and local revenue enhancement is not likely to be a long-term phenomenon. Within-state tax variation should be greater because of the increasing use of tax concessions and financial inducements. Between-state tax rate variations are also greater, as noted above. Thus, federal interference with interstate tax competition tends to create greater disparity in the taxes that citizens pay to state and local governments, even though they do not pay higher taxes on average *to these government units*, in the long run. This average includes those who receive special concessions, of course, so the "average" citizen who receives no tax concessions pays higher taxes even though average taxes still reflect interstate competition. Furthermore, *total* taxes paid by citizens are likely to be higher with federal revenue transfer programs because of "carrying charges," to use President Reagan's terminology (unless there really are significant economies of scale in tax collection).

Finally, federal programs have redirected the competitive incentives, leading politicians to innovate, and have developed many new forms of nontax-rate competition. Now that these mechanisms have been instituted, even if federal programs muting tax-rate competition are reduced, a rapid return to such competition is unlikely. Politicians now know that they can compete for new and mobile businesses without holding tax rates down for everyone. Vedder's (1990) findings that taxes matter more than ever in terms of their impact on economic growth is not surprising in this context, since tax-rate competition is less intense and many people pay higher taxes than ever as a consequence.

Conclusion

Competition in the political sector appears to have some of the same impacts that it has in the private sector: In particular, it constrains behavior. But as in the private sector, those who feel constrained have incentives to reduce competition. State governments have attempted to reduce interstate tax competition, and while apparently unsuccessful on their own, they do appear to use the offices of the federal government with some success. The existence of federally established institutions, such as deductibility of state and local taxes from federal taxes, mute the impact of interstate tax-rate competition.

federal programs rather than from long-established programs, so they have incentives to lobby for fairly continuous changes in the federal transfer process. Furthermore, it may help explain why states did not resist more strongly the relative reduction in federal transfers under the Reagan administration.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

However, the benefits that state and local politicians gain from reduced interstate tax competition are at least partially dissipated through nontax forms of competition. Incentives to compete do not disappear in the face of a lessening of the effectiveness of one particular avenue of competition.

References

Advisory Commission on Intergovernmental Relations (ACIR). *Regional Growth: Interstate Tax Competition.* Washington, D.C.: ACIR, 1981.

Benson, Bruce L. "Rent Seeking from a Property Rights Perspective." *Southern Economic Journal* 51 (October 1984): 388–400.

Benson, Bruce L., and Johnson, Ronald N. "Capital Formation and Interstate Tax Competition." In *Taxation and the Deficit Economy: Fiscal Policy and Capital Formation in the United States,* pp. 407–36. Edited by Dwight Lee. Cambridge, Mass.: Ballinger Press, 1986a.

Benson, Bruce L., and Johnson, Ronald N. "The Lagged Impact of State and Local Taxes on Economic Activity and Political Behavior." *Economic Inquiry* 24 (July 1986b): 389–401.

Benson, Bruce L., and Johnson, Ronald N. "Intergovernmental Competition and Efforts to Collude." Florida State University Working Paper, 1987.

Carlson, Eugene. "Great Lakes Governors Split over Truce on Industry Raids." *Wall Street Journal,* 12 July 1983, p. 31.

Church, Albert M. *Taxation of Nonrenewable Resources.* Lexington, Mass.: Lexington Books, 1981.

Douglas, G., and Miller, J. "Quality Competition, Industry Equilibrium, and Efficiency in the Price-Constrained Airline Market." *American Economic Review* 64 (September 1974): 657–69.

Friedman, David D., and Kurth, Michael M. "Revenue Sharing and Monopoly Government: A Comment." *Public Choice* 37 (1981): 365–70.

Fournier, Gary M. "Nonprice Competition and the Dissipation of Rents from Television Regulation." *Southern Economic Journal* (January 1985): 754–65.

Juster, Thomas, ed. *The Economic and Political Impact of General Revenue Sharing.* Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan, 1977.

Kleine, Robert J. "State and Local Tax Levels and Economic Growth—A Regional Comparison." *Proceedings of the Annual Conference of the National Tax Association—Tax Institute of America* 70 (November 1977): 162–72.

Kolstad, Charles D., and Wolak, Frank A., Jr. "Competition in Interregional Taxation: The Case of Western Coal." *Journal of Political Economy* 91 (June 1983): 443–60.

Lee, Dwight, and Buchanan, James M. "Tax Rates and Tax Revenues in Political Equilibrium: Some Simple Analytics." *Economic Inquiry* 20 (July 1982): 344–54.

McGuire, Martin C. "The Analysis of Federal Grants into Price and Income Components." In *Fiscal Federalism and Grants-in-Aid,* pp. 31–50. Edited

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

by P. Mieszkowski and W. Oakland. Washington D.C.: The Urban Institute, 1979.
McKenzie, Richard B., and Staff, Robert B. "Revenue Sharing and Monopoly Government." *Public Choice* 33 (1978): 93–97.
McKenzie, Richard B., and Staff, Robert B. "Revenue Sharing and Monopoly Government: A Reply." *Public Choice* 37 (1981): 371–75.
Miller, James P. "Interstate Competition for Business: Changing Roles of Federal and State Incentives." *Economic Research Staff Report,* No. AGES 8310112. Economic Development Division, U.S. Department of Agriculture, December 1983.
Posner, Richard. "The Social Cost of Monopoly and Regulation." *Journal of Political Economy* 83 (August 1975): 807–27.
Stigler, George. "Price and Non-Price Competition." *Journal of Political Economy* 76 (January 1968): 149–54.
Stigler, George. "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science* 2 (Spring 1971): 3–21.
Tollison, Robert D. "Rent Seeking: A Survey." *Kyklos* 35 (1982): 575–602.
Tullock, Gordon. "The Welfare Costs of Tariffs, Monopolies, and Theft." *Western Economic Journal* 5 (June 1967): 224–32.
Vedder, Richard. "Tiebout, Taxes, and Economic Growth." *Cato Journal* 10 (Spring/Summer 1990): 91–108.
Zimmerman, M. B., and Alt, C. "The Western Coal Cartel." Mimeo. Hoover Institute, Stanford University Working Paper E–80–9, March 1981.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TIEBOUT, TAXES, AND ECONOMIC GROWTH

Richard Vedder

Over the past decade or so, professional economists have increasingly questioned the once prevailing conventional wisdom that state and local taxes are unimportant in determining variations in the economic performance of states. While virtually no economist today takes the extreme view that "taxes alone matter" in determining state economic performance, few adhere to the once popular but equally extreme view that "taxes don't matter." A considerable body of research has demonstrated that taxes have an impact on economic performance, measured by growth in real per capita incomes, in the number of jobs created, or by statistics on the migration of human or capital resources (see Vedder 1989).

Given the growing realization that taxes matter, have states altered their tax structures in a way that reflects greater tax competition for jobs, capital, and income? Has interstate competition eroded the monopoly position of governments in providing public goods and services, which has led, perhaps, to some convergence in the tax/expenditure behaviors of states? Or has knowledge of interstate variations in providing and pricing public services led to migration that has accentuated these differences? Finally, have historical trends (whatever they are) worked to erode or enhance the observed impact that taxes have on economic growth at the state and local level? Do taxes matter as much as (or more than) they used to? These are some issues that this paper attempts to address.

Two Models of State and Local Provision of Services

There are different ways to conceptualize the relationship between state and local governments. Two different, but not necessarily

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at Ohio University.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

totally inconsistent, perspectives are outlined below and may help us examine the question of variations in state and local taxation and provision of services.

The Theory of the Firm/Public Choice Approach

The first perspective is a variant of the Leviathan model (Brennan and Buchanan 1980).[1] Consider that each state has one "firm," which we call state government, that provides certain goods and services. Local governments are legally subordinate to this single firm and may be viewed as subcontractors of it. While the services that the state government provides are diverse, for many (if not most) activities, the state government has a monopoly or near-monopoly on providing services within the market area. Moreover, unlike private firms, the governmental monopolist is a price-searcher who can compel "purchase" of services by levying taxes.

Taxes may be viewed as the price paid for the bundle of governmental services provided by the state-firm. While the state is a monopolist, the monopoly is not pure in that customers (taxpayers) have the option of moving to another jurisdiction (having public services provided by another governmental firm) if taxes become too onerous. Thus, interstate tax competition might force high-tax states to lower their prices (taxes), promoting tax convergence.

While this perspective on government has been developed as an analogy to the theory of the firm, added insight is possible from public choice economics. Governmental firms use tax revenues to provide two things: goods and services to their tax-paying customers, and income in the form of economic rents to certain individuals. The income may take the form of transfer payments (e.g., welfare benefits, pensions to public employees) or simply compensation to public employees beyond levels required by market forces.

Presumably as taxes rise, the real provision of governmental goods and services grows, but so do economic rents. Indeed, it seems intuitively plausible that when taxes and governmental services are at a "barebone" level, little tax money is available for economic rents. But as taxes rise, rent payments to welfare beneficiaries, governmental employees, etc., grow faster than taxes. Services also increase, but less rapidly than taxes. At a relatively high tax level, taxpayers revolt, because the price of government services is viewed as too high in relation to the volume or quality of services provided. The sheer

[1]The Leviathan model has been empirically supported by some scholars (Bell 1988). The theoretical and empirical literature on tax competition is substantial. Two quite different but representative samples are Beck (1983) and Wildasin (1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

quantity of services provided per dollar of incremental taxes falls as economic rents rise in relative, as well as absolute, terms.

While a full empirical testing of this hypothesis is beyond the scope of this study, some limited evidence is supportive. Specifically, from regression analysis using fiscal year 1988 data, we observe that elasticity of public welfare spending with respect to state and local taxes is estimated to be 1.21 for the 48 contiguous states.[2] Similarly, for the same states and year, average teacher salaries are estimated to rise $3.47 for each dollar of increased state and local taxes, even after we control for local labor market conditions. While this latter conclusion does not necessarily prove that the above hypothesis is valid, it is certainly consistent with the assumption.

This hypothesis is also consistent with the notion that there is some natural ceiling to the tax prices that are politically possible in each state jurisdiction. Moreover, given the migration of labor and capital, high-tax jurisdictions will feel economic pressure to lower the price of publicly provided services by offering tax cuts. Thus, the tax competition model suggests that, over time, tax disparities between high- and lower-tax jurisdictions may narrow in magnitude—there will be some tax convergence. Yet there never will be complete convergence. Aside from the fact that competition among the states may not be perfect, as long as there are some differences in tastes with respect to governmental services among states, there will be different quantities of governmental services provided, and thus some differences in tax prices and burdens.

The Tiebout Hypothesis

This last point leads us to the landmark model devised by Charles Tiebout (1956). Individuals have different tastes for governmental services, and thus they can shop around for the state that provides the menu of services most consistent with their tastes. State A may have relatively high taxes and high levels of governmental services, while State B has low taxes and services. Individuals with strong tastes for governmentally provided goods will locate in State A, whereas rugged individualists will move to State B. Thus the populations of A and B will diverge in terms of political philosophy, and those differences will be maintained and probably will even grow over time by migration: Frustrated conservatives will move from

[2]The log of income per capita is introduced as a control variable into the regression. Any elasticity greater than 1.0 is consistent with the hypothesis that the transfer-payment, nonservice component of state and local spending rises faster than the service component of such spending as taxes rise.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

State A to more congenial State B, whereas liberals will move from State B to State A.

If this perspective is correct, one might expect differences among states in tax regimes to be maintained and even to grow over time. Rather than serving as a force to bring convergence among states, migration would enhance interstate fiscal differentials. Even if migration does not serve to narrow the price differentials of government services per unit of services provided, it may well change the *quantity* of those services (and thus the overall tax burden) in a manner promoting tax divergence.

Federal Taxes and Their Impact

One other factor could have an impact on the degree of convergence or divergence over time, namely the deductibility of state and local taxes on federal income tax forms. When marginal tax rates are high, most of the individual taxpayer's costs of state and local taxes on income, property, and, until recently, sales are borne by the federal government when individuals itemize deductions. Thus the federal government provides an implicit subsidy for states to levy taxes. One unambiguous impact of this arrangement is that state and local governments have incentives to levy higher taxes than otherwise.

What this arrangement says for tax convergence, however, is not intuitively obvious. My a priori expectation is that tax deductibility promotes tax convergence. The reasoning here is that states with a tradition of high taxes and high levels of services face relatively little taxpayer resistance to increased governmental activity, so tax breaks are relatively unimportant in determining the size of governmental enterprise. States with a low-tax tradition, however, would normally resist expanded taxes. But when the subsidy from the federal government is sufficiently high, the resistance breaks down, leading low-tax states to raise taxes in a manner to reduce interstate tax differentials.

More technically, the elasticity of demand for governmental services with respect to taxes is hypothesized to be higher in conservative states than liberal ones. Tax deductibility lowers the tax price in all jurisdictions, but that lower tax rate increases the quantity of services demanded in conservative states relative to liberal ones, bringing about increased tax convergence.

A Brief History of Interstate Tax Differentials

What is the actual American experience? To answer this question, I compiled state and local tax revenue data for 1902, 1942, 1962,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

1970, 1980, and 1988. For the first three years, the data were obtained from special reports of the U.S. Census Bureau, whereas for the last three years, the data were taken from the annual publication *Governmental Finances in (Year)*. There are various aggregate tax measures. The one chosen here was "tax revenues" from "own sources." Thus federal grants are excluded, as well as state and local revenues from nontax sources, including fees, charges, and business operations. Data are reported on a per capita basis and per $1,000 of personal income. The per capita measure is the best indicator of the absolute level of governmental tax activity, whereas the data related to personal income are a better measure of the true tax burden or the aggregate tax rate in each jurisdiction. A very low-income state can have fairly low per capita levels of taxation, yet the average (and marginal) rates of taxation may be relatively high.

Table 1 shows the trend in the mean magnitudes of taxation over time, as well as the coefficient of variation. Note that by any standard, state and local taxes rose intertemporally until 1970. Since 1970, however, the average aggregate state and local tax rate (as measured by taxes per $1,000 of personal income) has reached a plateau and even declined slightly.

Regarding tax convergence/divergence, note that from 1902 to 1942, taxes converged dramatically when measured on a per capita basis, but they actually diverged somewhat when measured in rela-

TABLE 1

STATE AND LOCAL TAX LEVELS AND VARIATIONS IN THE UNITED STATES, 1902–88

	Per Capita Basis		Per $1,000 Personal Income	
Year	Average[a]	Coefficient of Variation	Average	Coefficient of Variation
1902	$ 110.81	.4851	$ 46.30	.1829
1942	368.29	.2877	69.86	.2018
1953	484.68	.2340	78.30	.1644
1962	695.46	.2062	95.30	.1130
1970	1,011.78	.2126	115.41	.1230
1980	1,116.64	.2046	112.45	.1323
1988	1,386.82	.2151	113.66	.1249

[a]In constant 1982–84 dollars, deflated by the Consumer Price Index. Means are unweighted arithmetic means of observations for the 48 contiguous states.

SOURCE: Author's calculation is from U.S. Bureau of the Census data for various years.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tion to personal income levels. The reason for this dichotomy is that there was substantial convergence in interstate per capita income differentials during this era. There is unambiguous convergence in state and local tax burdens in the 1942 to 1962 period. The convergence is pronounced using either the per capita or personal income measure. After 1962, however, convergence stops by either measure. Using either the per capita or personal income measure, we see that interstate tax differentials *increase* slightly from 1962 to 1988, although the move to greater diversity is not consistent within that period. The safest interpretation of the findings is probably that the move toward convergence observed before 1962 has halted, and there has been rough stability in the magnitude of interstate tax differentials since.

While the 1902–42 evidence might be regarded as mixed, the 1942–62 period appears characterized by tax convergence that seems consistent with the initial model of price (tax) competition among the various state governments. The post-1962 evidence is consistent with a view that some equilibrium amount of tax diversity exists. Some disparity among taxes will always exist in equilibrium, in part because of the Tiebout hypothesis, in part because of the costs of resource movement to eliminate tax disparities.

More generally, the changing patterns of tax convergence seem to be closely tied to changes in federal income taxation. Between 1902 and 1942, there was substantial diversity among states in tax systems, a diversity that was maintained throughout the period (particularly if one uses the personal income measure of tax burden). It was also a period of very limited federal income taxation, with most people not subject to taxation (indeed, in the early part of the period, there was *no* income tax.) In the absence of widespread federal income taxation, states maintained great diversity.

From 1942 to 1962, tax systems converged. At the same time, the federal income tax became universal (meaning most families were subject to it) for the first time; marginal tax rates were high and, partly because of inflation, rising. Federal income tax deductibility became a significant factor; the marginal cost to taxpayers of increased state and local taxes was reduced by federal "tax expenditures" in the form of state and local tax deductions against federal taxable income.

In the quarter century from 1962 to 1987, the observed reduction in tax diversity was halted, at a time when there were significant efforts to roll back the debilitating high marginal taxes imposed during the 1930s and, especially, during World War II. Although bracket creep from continuing inflation tended to push marginal tax rates up, major tax cuts in 1964–65, 1981, and 1986 served to reduce

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the value of federal income tax deductions by the end of the period. Thus the marginal cost to the taxpayer of a given state or local tax increase began to rise again. This rise increased taxpayer resistance in states, particularly those with relatively conservative political traditions (an elastic demand for government services).

This observation is verified by looking at the changing tax behavior of high- and low-tax states over time (Table 2). I divided the 48 contiguous states into quartiles according to the amount of taxes levied per $1,000 of personal income. In 1962, the quartile of states with the lowest tax burdens had increased that tax burden by more than $24 per $1,000 of income over the previous 20 years. This figure was nearly as much as the average tax increase in the quartile of states with the highest tax burden in 1962. Indeed, in percentage terms, the low-tax states raised their taxes *more* than the high-tax ones during the previous 20 years. In an era of high federal marginal tax rates, historically low-tax states responded to the powerful incentives to raise taxes provided by income tax deductibility.

By contrast, in 1987, the lowest quartile of states in terms of overall tax burden had raised their taxes during 1962–87 by an average of only $4.34 per $1,000 of personal income. This behavior was in marked contrast to the high-tax states, which raised their taxes by an average of $27.72 per $1,000 of personal income.

In short, during the 1940s and 1950s, all states raised their taxes substantially, the low-tax states about as much as the high-tax ones,

TABLE 2

CHANGING TAX BEHAVIOR, HIGH- AND LOW-TAX STATES, 1962 AND 1987

	Change in Taxes over Previous Period[a]		Percent Change in Taxes over Previous Period[a]	
Quartile[b]	1962	1987	1962	1987
Lowest	$24.07	$ 4.34	41.53%	4.71%
Second	21.88	14.87	32.11	16.24
Third	25.75	21.28	34.64	22.84
Highest	28.89	27.72	36.57	26.85

[a]For 1962, the previous period is from 1942 to 1962; for 1987, the previous period is from 1962 to 1987. Taxes are total raised from own sources per $1,000 of personal income; numbers are unweighted means of the 12 observations.

[b]The 48 contiguous states are divided into four groups of 12 states each. Lowest refers to the group that in the year indicated had the lowest tax burden per $1,000 of personal income.

SOURCE: Author's calculations using U.S. Bureau of the Census data.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

bringing about a reduction in the *relative* tax differential among states. After 1962, some low-tax states seemed to reach a political threshold beyond which they would not raise taxes much further, while some high-tax states continued to increase taxes as before, increasing tax divergence. An extreme example is Mississippi, which moved from the top quartile in terms of tax burden in 1962 to the bottom quartile by 1987, simply by reducing that burden modestly in the intervening years (when most states were still raising taxes somewhat). This post-1962 behavior would seem highly consistent with the Tiebout hypothesis. People in conservative states (where the elasticity of demand for governmental services was highly elastic with respect to price) seemed to engage in tax revolt strategies, whereas in liberal states tax-financed expansion in governmental services still seemed to garner political support.

A simple model was developed to provide further insight into the convergence and competition issue. The model empirically examined the 48 contiguous states and the District of Columbia (49 observations):

$$(1)\ CHTAX = a + b\ 67TAX + c\ 67INCAP + d\ \%GOP + e\ NEIGHTAX + f\ CHGNEIGHTAX + u,$$

where *CHTAX* is the change in the aggregate state and local tax rate from 1967 to 1987, as measured by tax revenues per $1,000 of personal income; *67TAX* is the tax rate in the initial year 1967; *67INCPAP* is the state's 1967 per capita income level; *%GOP* is an indicator of the conservativeness of a state, as measured by the average percentage of people voting for the Republican candidate in the 1976 and 1988 presidential elections; *NEIGHTAX* is the average aggregate tax rate in 1967 of all states bordering on the state in question; and *CHGNEIGHTAX* is the change from 1967 to 1987 in the aggregate tax rate in the bordering states. The lowercase letters represent the constant, regression coefficients, and random error term.[3]

If convergence is occurring, one might expect a statistically significant negative relationship between *67TAX* and *CHTAX*—high-tax states would raise taxes less than low-tax ones. If states were concerned about tax prices of other states, presumably that concern

[3]The model used 1967 as the initial year for two reasons. First, most of the large national increase in aggregate state and local tax rates had been completed by that year, so 1967–87 was a period of relative stability in state and local tax average rates (compared with 1962–87, since taxes were still rising rapidly in the early and mid-1960s). Second, a major change in federal income tax laws took place in 1964 and 1965, lowering marginal tax rates. Because of the potential importance of those rates, the analysis begins after the 1964–65 tax cut was fully implemented.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

would be greatest with respect to neighboring states. Arizona, for example, might want to keep taxes low to lure the XYZ Corporation there as opposed to its being lured to some other state in the Southwest, such as neighboring California, Nevada, or Utah. If sensitivity to competitive tax strategies were important, presumably there would be a significant positive relationship between *NEIGHTAX* and *CHTAX*, and also between *CHGNEIGHTAX* and *CHTAX*. If neighbors have initially high and/or rapidly rising tax rates, then one can afford to raise taxes more than if neighbors have initially low and/or falling tax rates.

The previous analysis indicated that in periods of a decline in the federal marginal tax rate, such as occurred during the last part of 1967–87, we might expect politically conservative areas to become more resistant to tax hikes, as the cost of those hikes grow. Thus it is hypothesized that *%GOP* and *CHTAX* are negatively related. *INCP*67 is introduced for control purposes with no sign postulated a priori.

The results of estimating equation (1) by ordinary least squares regression procedures are indicated in Table 3. Essentially, the key tax and tax competition variables are all statistically insignificant and one has the wrong sign. There is no evidence supporting the hypothesis of tax convergence on the basis of the response to tax behavior of neighboring states. There is, however, evidence that suggests relatively conservative states received more of an impact from political opposition to tax increases and from a tax revolt phenomenon than did relatively liberal states. The mean change in state and local taxes per $1,000 of income from 1967 to 1987 for the most

TABLE 3

REGRESSION RESULTS: U.S. TAX CONVERGENCE MODEL, 1967–87

Variable or Statistic	Coefficient or Statistic	T-Value
Constant	53.065	1.990
67*TAX*	−0.161	0.949
67*INCP*	0.014	2.866[a]
%GOP	−0.664	2.179[b]
NEIGHTAX	−0.312	1.073
CHGNEIGHTAX	0.026	0.092
$\bar{R}^2$	.391	—
F-Statistic	7.151	—

[a]Statistically significant at the 1 percent level.
[b]Statistically significant at the 5 percent level.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

conservative quartile of states was a negligible $0.76; for the most liberal quartile of states, the mean change was $9.16, or 12 times as much.[4]

It is interesting, other factors held equal, that the aggregate tax rate was likely to rise more the higher the level of per capita income in 1967. Rich states (in the beginning year) were more willing to tax themselves more over time to support government. Whether that behavior contributed to a loss of economic advantage for those prosperous states is discussed shortly.[5]

The Role of Nontax Sources of Revenue

The discussion above has focused on tax revenues raised by state and local governments. Yet those governments have two other sources of income: federal grants and nontax sources of revenues such as fees, user charges, and interest income. Since 1967, the nontax forms of revenue have become relatively more important in financing state and local governmental activities (see Table 4).

TABLE 4

CHANGES IN STATE AND LOCAL GOVERNMENTAL REVENUES IN THE UNITED STATES, 1967–88, PER $1,000 OF PERSONAL INCOME

Sources of Funds	Levels		Absolute Change	Percent Change
	1967	1988		
Taxes	$106.13	$113.66	$ 7.53	7.1%
Federal Aid	33.60	35.07	1.47	4.4
Nontax Revenues	30.00	50.88	20.88	69.6
All Revenues	169.73	199.61	29.88	17.6

SOURCE: Author's calculations using data from *Governmental Finances in 1967, Governmental Finances in 1988.*

[4]Liberal and conservative states are defined, as before, by the average percentage of people voting for the Republican candidate in the 1976 and 1988 elections. The District of Columbia is excluded in the calculations. If we include D.C., the conservative to liberal comparison is much greater.

[5]Two topics ignored in this discussion of tax convergence are (1) exporting tax burdens and (2) tax abatement and other development initiatives. States with large oil exporting capabilities such as Alaska have been able to export much of their tax burden, promoting tax divergence. Also, there has been a major increase in tax competition for business, which I call micro-tax competition. Authorities differ in their assessment of the economic impact of micro-tax competition. Some seem to believe that on balance it enforces a discipline of public officials that is healthy (Browning and Browning 1987). Other studies have suggested that when there are positive rates of capital taxation, tax competition of this sort can have distortive effects (Oates and Schwab 1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Indeed, nearly 70 percent of the growth in the relative size of government receipts is explainable by the explosive growth in nontax revenues.

Taxes and federal grants seem to be largely formulated according to the "ability to pay" principle of public finance, while nontax revenues such as fees and interest income reflect (to a considerable extent) the "benefit principle." Thus it appears there has been some shift toward the benefit principle in financing state and local government.

Why the shift? The following explanation deserves consideration. First, the ability of state and local governments to raise taxes has been severely restricted. As tax rates have risen over time, so has taxpayer resistance. Thus the elasticity of demand for comprehensively financed government services has also risen. This rise is particularly true in relatively conservative states, where the decline in the federal tax break because of falling federal marginal tax rates has increased state and local tax prices of services and has led to taxpayer revolts.

With tax revenues closed off and with a disinclination for the federal government to expand funding for the states, advocates of increased state and local governmental activity have been forced to use nontax means of expenditure, particularly in the more conservative states. This shift is confirmed empirically by regressing the proportion of increased tax revenues financed by nontax means over the 1967–88 period against the conservativeness of a state (as measured by the average percentage of people voting for Republican candidates in the 1976 and 1988 elections) and the average tax rate in 1967. There is a statistically significant positive relationship between a state's conservatism and its increased reliance on nontax revenues; also higher-tax states in 1967 relied more on nontax sources to increase revenues after that date.

Taxes and Economic Growth: Contemporary Studies

If we turn to the economic impact of tax policy until the 1970s (some would say the 1980s, we observe the conventional wisdom that, in general, state and local tax/expenditure policies had only a minor impact on economic conditions. Looking specifically at business location, for example, John Due (1961) concluded that studies "suggest very strongly that the tax effects cannot be of major importance." Even in the late 1970s, some scholars continued to reach a similar conclusion (Oakland 1978).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The 1970s, however, were a turning point; increasingly, scholars questioned the view that taxes do not matter much. That conclusion was questioned in three different ways. First, a spate of new studies at the national level suggested that federal taxes had debilitating effects on various ingredients of economic progress, notably capital formation. These studies suggested, for example, that capital taxation tends to retard savings and investment, and that Social Security taxation similarly retards savings. Also, taxes or expenditures often have unintended consequences: For example, taxes on labor income could also repel capital.[6]

A second group of studies focused specifically on state and local tax policy. The view that taxes are irrelevant to business location decisions was questioned in a major econometric study of that topic (Grieson, Hamovitch, and Morgenstern 1977). Other studies opined that taxes repel in-migration (Cebula 1974, Browne 1979). Finally, simple econometric studies of the tax-growth relationship began to appear and concluded that state and local taxes have an adverse impact on economic growth (Genetski and Chin 1978, Vedder 1981).

A third force behind the view that taxes matter a good deal came from various popularizers of supply-side economics. While Arthur Laffer was much scorned by many, his Laffer curve ingeniously drove home the point that severe tax disincentive effects could lead to unintended revenue consequences. The editorial page of the *Wall Street Journal* articulated a "supply-side" perspective. One journalist, Warren Brookes, stands out as being particularly effective in offering simple empirical evidence of the negative impact of state and local taxes (see Brookes 1982). While Howard Jarvis and Paul Gann are known as the starters of the modern tax revolt in California in 1978 with Proposition 13, a variety of citizens' groups, aided by national organizations like the National Tax Limitation Committee and the National Taxpayers Union, increased agitation for tax reduction in many states. In Washington, economists such as Paul Craig Roberts and Norman Ture, as well as the Joint Economic Committee, worked to sell the view that taxes do matter.

As the 1980s have progressed, further research has refined but essentially confirmed the conclusions of earlier scholars. A few examples are worth noting. Jay Helms (1985) used more sophisticated econometric techniques than earlier scholars, and he related tax changes to the uses made of new tax revenues. He concluded that, in general, there was a negative relationship between taxes and

[6]On this point, see McLure (1970), Feldstein (1976), and Boskin (1978); on expenditures, see Hall (1980).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

economic growth, and that the negative relationship was clearly pervasive when tax increases are used to finance transfer payments. However, Helms implied that the tax-growth relationship might not be negative if new tax revenues went into more growth-inducing activities such as education or highways.

Using a distributed-lag regression model, Bruce Benson and Ronald Johnson (1986) found that taxes impeded economic growth, but that much of the impact comes two or more years after the initial period in which the tax takes effect, suggesting that politicians can often escape the negative consequences of their actions as they reap some political gains from immediate increases in spending. In short, there is a shortsightedness effect associated with state and local tax increases.

Finally, new studies on business location or the birth of new business ventures reconfirm the importance of taxes. Timothy Bartick (1987) found marginal personal income tax rates inversely related to small business start-ups. A recent study by Paul Bauer and Brian Cromwell (1989) examining variations in new firm creation in more than 250 metropolitan areas shows a highly significant negative relationship between the effective corporate tax rate and business creation. The model includes nearly 20 other independent variables (mostly relating to characteristics of local banking institutions).

Intertemporal Changes in the Tax-Growth Relation

As indicated above, the modern view that taxes do matter evolved largely in the late 1970s and early 1980s. The empirical evidence to support this view was largely derived from the growth experience of the 1970s. Yet the 1980s have seen a dramatic change in the relative growth experience of states, as can be seen in Table 5. During the

TABLE 5

CORRELATION BETWEEN GROWTH IN DECADE AND PREVIOUS DECADE

Decade	Correlation of Growth Rate with Previous Decade[a]	Statistically Significant?[b]
1940s[c]	−0.0749	No
1950s	+0.1332	No
1960s	+0.4572	Yes
1970s	+0.1536	No
1980s	−0.5913	Yes

[a]For 48 contiguous states.
[b]At the 1 percent level.
[c]Growth for the 1940s is related to growth during 1929–40.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

four decades prior to the 1980s, the correlation between economic growth rates and economic growth rates in the preceding decade was either positive or not significantly different from zero. States that grew substantially during the 1950s, for example, tended to grow by relatively large amounts during the 1960s. Yet when one compares the states' growth rates for the 1970s with the rates for the 1980s, one obtains a strongly *negative* correlation: Fast-growing states in the 1970s have typically been relatively slow-growing states in the 1980s, and vice versa. This significant and large negative correlation is not only unusual—it seems to fly in the face of a long-run trend that is highly consistent with economic theory, the trend of factor price equalization. With the help of the migration of human and physical capital, interregional income differences have tended to narrow over time. Yet in the 1980s, high-income states like Massachusetts and New York grew faster than low-income states like West Virginia and Mississippi.

Thus it is possible that observed negative associations between new tax initiatives and growth in the 1970s would not be repeated in the 1980s. Nontax factors may explain growth patterns in the 1980s, just as tax factors were important in explaining them in the 1970s. We are told, for example, that America's unionized Rust Belt has priced itself out of world markets. Also, falling relative energy prices have had an impact on numerous energy-producing states that boomed from rising relative energy prices in the 1970s.

Accordingly, I have developed a simple model attempting to explain the percentage of variations in real per capita personal income in two periods: 1970–80, and 1980–88. As the tax variable, *CHGTAX*, I used the growth in state and local taxes per $1,000 of personal income over the relevant time period. Four other independent variables were introduced for control purposes: (1) the level of real per capita income at the beginning of the relevant time period, designated *INCOME*; (2) the proportion of the labor force belonging to labor unions at the beginning of the period, labeled *UNION*; (3) the age of the state as measured by years since statehood, designated *STATEAGE*; and (4) the proportion of personal income derived from mineral (predominantly fuel) production in 1980, called *ENERGY80*.[7]

[7]The use of years since statehood is suggested by Olson (1982). The variable measuring intensity of energy production is calculated as of 1980 for both periods. Because of new energy discoveries and rising relative prices, energy production rose dramatically in some states in the 1970s, so the end-of-period data are used in the 1970–80 growth model.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The results are indicated in Table 6. The expected negative relationship between changes in the aggregate tax rate and economic growth is observed in both periods. The relationship is actually stronger in the 1980s than the 1970s. A $10 increase in taxes out of every $1,000 in personal income (a 1 percent share) is estimated to have reduced real growth by 0.72 percent during 1970–80, but by 2.30 percent during 1980–87. Since the mean growth in the latter period was only 11.6 percent, the results suggest tax policy could have a material impact on the observed growth experience. As federal marginal tax rates fell in the 1980s and the implicit subsidy to taxpayers for state and local tax payments fell, the observed relationship between changes in tax policy and economic growth seemed to grow.

Note that the *ENERGY80*, *INCOME*, and *STATEAGE* variables reverse signs between the two decades, and the *UNION* variable moves from being insignificant to being statistically and significantly negative. Factors working to increase incomes in the 1970s worked to reduce them in the 1980s, or vice versa, except for tax changes. The continued negative correlation of tax changes to economic growth despite dramatic changes in other observed relationships enhances confidence in the enduring nature of the tax-growth relationship.

TABLE 6

DETERMINANTS OF VARIATIONS IN INTERSTATE ECONOMIC GROWTH: REGRESSION RESULTS FOR THE UNITED STATES, 1970–80 AND 1980–88

Statistic or Variable	1970–80[a]	1980–88[a]
Constant	31.076	−11.308
	(5.986)	(1.814)
CHGTAX	−0.072	−0.230
	(2.336)	(4.917)
INCOME	−0.001	0.002
	(2.005)	(3.043)
ENERGY80	0.251	−0.301
	(5.501)	(6.538)
UNION	−0.077	−0.292
	(0.969)	(2.754)
STATEAGE	−0.004	0.113
	(0.236)	(5.978)
$\bar{R}^2$	.625	.737
F-Statistic	17.631	29.012

[a]*Numbers in parentheses are t-statistics.*

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Actually, the negative tax-growth relationship was well established even before economists were almost unaminously agreeing that taxes do not matter much (roughly 1950–75). For example, using economic growth data for 48 states for 1900–42, I regressed the economic growth rate against the initial tax rate (*TAX*) in 1900 (actually 1902), the change in tax rates from 1902 to 1942 (*CHGTAX*), and the per capita income levels in 1900 (*INCOME*):

$$(2)\ GROWTH = \underset{}{377.659} - \underset{(2.258)}{1.422\ TAX} - \underset{(4.917)}{1.967\ CHGTAX} - \underset{(9.323)}{1.967\ INCOME},\ \bar{R}^2 = .717,\ F = 40.702,$$

(11.461)

where the numbers in parentheses are t-values. In the era before modern computers, economists seemed to be ignorant (or ignored) a strong negative relationship between the levels and changes in taxes and the rate of economic growth.

Conclusion

It has been observed that Americans of different regions or states have become more similar over time: Income differentials have declined just as have regional differences in dialect. Falling communications and transactions costs have narrowed spatial distinctions. With respect to state and local tax behavior, a similar convergence has been observed over part of this century (e.g., 1942–62). Since 1962, however, convergence has stopped, and rough stability in tax differentials has been achieved. Tax competition and rising elasticity of demand for government services with respect to the tax price (partly reflecting falling federal income tax marginal rates) have made it harder for states to raise taxes. Nontax revenues have thus become relatively more important as more governmental activities are being financed by direct beneficiaries rather than by the general taxpayer. The Tiebout notions of differential political tastes probably have also contributed to the persistence of some tax differentials among states. The empirical evidence continues to suggest that the growth-inducing effects of governmental expenditures, on balance, are less than the growth-impeding effects of taxes used to finance those expenditures. Recent reductions in federal marginal tax rates should work to increase the intensity of the tax-growth relationship, and some limited empirical evidence presented here supports that theory. As sensitivity to taxes increases, it appears that state and local taxes are

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

becoming more, not less, important a determinant of variations in economic growth among the states.[8] Taxes matter more than ever.

References

Bartik, Timothy J. "The Effects of State and Local Government Policies on the Start-Up of Small Business." *1986 Proceedings of the 79th Annual Conference on Taxation.* Columbus, Ohio: National Tax Association, 1987, pp. 25–33.

Bauer, Paul W., and Cromwell, Brian A. "The Effect of Bank Structure and Profitability on Firm Openings." *Economic Review*, Federal Reserve Bank of Cleveland 25 (Fall 1989): 29–39.

Beck, J. H. "Tax Competition, Uniform Assessment, and the Benefit Principle." *Journal of Urban Economics* 13 (March 1983): 127–46.

Bell, Christopher R. "The Assignment of Fiscal Responsibility in a Federal State: An Empirical Assessment." *National Tax Journal* 41 (June 1988): 191–207.

Benson, Bruce, and Johnson, Ronald. "The Lagged Impact of State and Local Taxes on Economic Activity and Political Behavior." *Economic Inquiry* 24 (July 1986): 389–401.

Boskin, Michael. "Taxation, Savings, and the Rate of Interest." *Journal of Political Economy* 86 (April 1978): S3–27.

Brennan, Geoffrey, and Buchanan, James. *The Power to Tax.* Cambridge: Cambridge University Press, 1980.

Brookes, Warren. *The Economy in Mind.* New York: Universe Books, 1982.

Browne, Lynn. "The Shifting Pattern of Interregional Migration." *New England Economic Review* (November/December 1979): 17–32.

Browning, Edgar K., and Browning, Jacquelene M. *Public Finance and the Price System.* New York: Macmillan, 1987.

Cebula, Richard. "Local Government Policies and Migration: An Analysis for SMSAs in the United States, 1965–1970." *Public Choice* 19 (Fall 1974): 85–93.

Due, John. "Studies of State-Local Tax Influences on Location of Industry." *National Tax Journal* 14 (June 1961): 163–73.

Feldstein, Martin. "Social Security and Saving: The Extended Life Cycle Theory." *American Economic Review* 66 (May 1976): 77–86.

Genetski, Robert J., and Chin, Young D. "The Impact of State and Local Taxes on Economic Growth." Chicago: Harris Bank, 3 November 1978.

Grieson, Ronald; Hamovitch, William; and Morgenstern, Richard. "The Effects of Business Taxation on Industry." *Journal of Urban Economics* 4 (April 1977): 170–85.

[8]Using the regression results from Table 4, a $10 increase in state and local taxes in the 1970s would have lowered the growth on average by 0.72 percentage points, or 3.4 percent of the mean growth rate of 21.1 percent. In the 1980s, by contrast, a similar tax increase would have, on average, lowered growth by 2 percentage points, or 19.9 percent of the mean growth rate of 11.6 percent. If we use data for the 48 contiguous states, similar results are obtained.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hall, Robert. "Stabilization Policy and Capital Formation." *American Economic Review* 70 (May 1980): 157–63.

Helms, L. Jay. "The Effect of State and Local Taxes on Economic Growth: A Time Series-Cross Section Approach." *Review of Economics and Statistics* 67 (November 1985): 574–82.

McLure, Charles, Jr. "Taxation, Substitution, and Industrial Location." *Journal of Political Economy* 78 (January/February 1970): 112–32.

Oakland, W. H. "Local Taxes and Intra-urban Industrial Location: A Survey." In *Metro Financing and Growth Management Policies*. Edited by George F. Break. Madison: University of Wisconsin Press, 1978.

Oates, Wallace E., and Schwab, Robert M. "Economic Competition among Jurisdictions: Efficiency Enhancing or Distortion Inducing?" *Journal of Public Economics* 35 (April 1988): 333–54.

Olson, Mancur. *The Rise and Decline of Nations*. New Haven: Yale University Press, 1982.

Tiebout, Charles. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (October 1956): 416–24.

Vedder, Richard. "State and Local Economic Development Strategy: A 'Supply Side' Perspective." Staff Study, Joint Economic Committee of Congress. Washington: Government Printing Office, 1981.

Vedder, Richard. "Do Tax Increases Harm Economic Growth and Development?" Flagstaff, Ariz.: Barry Goldwater Institute Arizona Issue Analysis no. 106, 1989.

Wildasin, David E. "Nash Equilibrium in Models of Fiscal Competition." *Journal of Public Economics* 35 (March 1988): 229–40.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Growth Management in Florida: Lessons for the National Economy

Randall G. Holcombe

Introduction

In 1985, Florida's Growth Management Act was passed, and this paper discusses its characteristics and implications. The object of Florida's Growth Management Act is to take some of the rights to determine the use of land away from the property owner and to have the use of the land determined politically. The Act also transfers to the state government some land use decisions that used to be made by local governments through zoning. While the specific implications of the Act apply specifically to Florida, there are broader lessons for the nation. Growth management on a statewide basis has become increasingly popular over the past 15 years: A number of states have enacted statewide land use planning of some type or another, and other states are considering such legislation. Florida's experience is relevant to the nation because the reasons behind statewide growth management are the same in Florida as in other states, and because Florida's Growth Management Act could be used as a model for other states, much as Florida adopted many of the features of Oregon's growth management legislation. This paper considers both the general motivations for growth management and the way that growth management has been implemented in Florida.

Florida's Growth Management Act specifies a political process through which land use decisions are made. The details of the process are described below, but the effect of the Act is to take some of the right to determine how land will be used away from the individ-

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at Florida State University. He gratefully acknowledges helpful comments and background information from James Gwartney, John Metcalf, and David Rasmussen. The usual caveat applies.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ual property owner and to transfer it into the political arena. Clearly, individuals who do not own much land but do have political power will benefit from this transfer of rights away from landowners. This transfer provides a base of political support for statewide growth management. Growth management lets some people decide how the property of others will be used.

There is a clear redistributional aspect to growth management legislation that explains its popularity.[1] Aside from the distributional aspect, does growth management legislation improve land use patterns? For growth management at the state level, the answer is clearly no. Economists since Adam Smith have observed that private ownership and market allocation of resources are more efficient than government allocation, and the contemporary economic analysis of property rights supports this view.[2] On the other side of the argument, externalities may exist so that one person's land use decisions impose costs on everyone else. Bernard Siegan (1970, 1972) has explained how market allocation of land minimizes these externalities, but Siegan's argument is unnecessary for analysis of land use planning at the state level. Few land use externalities will spill over from one county to another, so local government planning should be sufficient to internalize any externalities arising from land use decisions.

Statewide growth management is undesirable because it creates more poorly defined property rights, which reduce the efficiency of land use decisions. Such management cannot be justified on the basis of externalities. Externalities will largely be confined to local government jurisdictions, such as counties; local governments already use tools such as zoning and eminent domain to deal with land use externalities.

The foregoing discussion illustrates the undesirability of statewide growth management in terms of general economic principles. The remainder of the paper examines Florida's Growth Management Act specifically to show how the Act has affected private property rights of landowners and how land use in Florida is likely to be affected as a result of Florida's specific implementation of growth management.

An Overview of Florida's Growth Management Act

Florida's Local Government Comprehensive Planning and Land Development Regulation Act (Florida's Growth Management Act)

[1]See Holcombe (1990) for a more complete discussion of the distributional aspects of Florida's Growth Management Act.

[2]See, for example, Alchian (1965) for a good discussion of the inefficiencies of governmentally determined resource allocation when compared to market allocation with private property rights.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

was passed in 1985, and its full effects have yet to be seen. In brief, the Act requires local governments to submit comprehensive land use plans to the Florida Department of Community Affairs. These plans must comply with the provisions of the Act. Two notable features of the Act are, first, that local plans put all land into zones that allow varying levels of development and, second, that there is a requirement that infrastructure be in place concurrent with development.

Florida has seen decades of rapid growth and has suffered some growing pains in the process. Florida's Growth Management Act is a response to those growing pains. The Act has widespread support, partly because people tend to think that when problems arise, the government should do something, and partly because the Act is relatively new and its negative consequences have not yet been felt. Despite the inevitable costs of the Act, it is not designed in such a way that it will succeed in achieving its stated goals. Unfortunately, for most of Florida's citizens who will not study the issue in depth, the negative consequences of the Act will appear to be the result of growth itself, which could lead to a call for additional legislation to address the effects of statewide growth management.

Some of the infrastructure-related growing pains in Florida have been the result of the fact that infrastructure lasts a long time, but it must be financed as it is built. Many of the future taxpayers who will use the infrastructure for decades to come do not yet live in the state and are not yet paying taxes. These financial difficulties are a product of growth itself that statewide comprehensive planning cannot hope to overcome. Some infrastructure-related growing pains are also a result of the lack of foresight of government planners decades ago. If more land were obtained for rights-of-way, parks, and other municipal uses when such land was in the path of future development, infrastructure-related problems in many urban areas would be less pressing.

It is difficult to find fault with government planners in the past for not being able to foresee the future, but current growth management legislation is written as if current planners can foresee the future. With hindsight, problems are apparent with regard to both government plans for growth and developers' plans for their own property. The real question for Florida today is whether Florida's Growth Management Act will be beneficial in guiding future growth and development. This paper argues that the Act will, on net, be detrimental to Florida.

Some Details of Florida's Growth Management Act

Florida's Growth Management Act is built around the comprehensive plan that must be submitted by all local governments and

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

approved by the state. Development is allowed only if the land use does not violate the local plan.[3] Because state approval of the local comprehensive plan is required, the Act gives the state extensive control over local land use in Florida.[4] Once drawn up, the local plans can be modified, which provides flexibility to the planning process. Just as with the original plans, any modifications must be in accordance with the other provisions of the Act.[5]

In this regard, perhaps the most significant provision of Florida's Growth Management Act is the so-called concurrency requirement. The concurrency rule requires "that public facilities and services needed to support development shall be available concurrent with the impacts of such development."[6] Stated this way, the concurrency goal is not controversial. Everyone wants the facilities needed to support development to be available concurrent with the development. Nevertheless, the concurrency requirement has become controversial because political implementation of concurrency means that development decisions that previously could be made by landowners can now be made only as part of a political decisionmaking process.

The importance of the concurrency requirement becomes clear when it is seen how concurrency is used to develop local comprehensive plans and to manage growth. The immediate effects of concurrency are most binding on roads, so the following discussion will concern that aspect of the infrastructure. Note, however, that the concurrency requirement applies to other infrastructure components

[3]To promote orderly development, the Department of Community Affairs encourages the local comprehensive plans to conform to population projections for that area. For example, the Department of Community Affairs would question a plan that would accommodate twice the population projected for the area. As a rule of thumb, the Department of Community Affairs does not want comprehensive plans to accommodate more than 120 percent of an area's projected population.

[4]In testimony before the Florida House Committee on Community Affairs on November 13, 1989, Thomas Pelham, Secretary of the Department of Community Affairs, remarked that one county submitted a local plan to the state that had zoned all land adjacent to state roads for commercial use. Secretary Pelham viewed this as undesirable because it could create "urban sprawl," and the plan was rejected by the state. This example clearly shows the extent of state control in decisions that previously would be under the jurisdiction of local governments.

[5]Florida's Growth Management Act allows modifications only twice a year, which limits flexibility. A larger factor limiting flexibility might be that there usually has been a sufficient supply of land zoned for the type of development that developers were undertaking. Local comprehensive plans could limit the supply of land zoned in certain ways so as to make zoning itself more binding, independent of the other provisions of the Act.

[6]This wording is quoted from Florida Statute 163.3177(9).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

as well, including wastewater, schools, parks, and fire and police protection.

The local comprehensive plans must establish level-of-service standards for all roads within the local government's jurisdiction. There are legal definitions both for levels of service, and for levels of service that will be acceptable in the comprehensive plans. If existing levels of service are unacceptable, then the state will not approve a comprehensive plan that allows additional development that will add to the traffic on an unacceptably congested road. To be approved, a comprehensive plan must provide for a way to improve levels of service to acceptable levels.[7]

The first political hurdle regarding concurrency should now be apparent. Owners of undeveloped property can develop that property only if development is consistent with the local comprehensive plan. The comprehensive plan can incorporate development only if it does not degrade levels of service to unacceptable levels. Therefore, for those interested in immediate development, much is at stake in having that development included in the comprehensive plan.

The concurrency requirement has the potential for even greater effects on future development. For future development to be approved, infrastructure must be in place concurrent with the development. Concurrency in roads, for example, could be especially troublesome to developers because local comprehensive plans define levels of service on roads. If a development will add traffic to a road several miles away (closer toward downtown, for example) so that the road would decline to a level of service below that defined in the comprehensive plan, the development would violate the concurrency rule and so would not be allowed.

The potential impact of the concurrency rule is heightened because, as will be discussed further below, almost anyone can challenge a development on concurrency grounds.[8] Thus, Florida's Growth Management Act gives people who are opposed to a development for any reason a legal mechanism for stalling the development and for imposing costs on the developer.

[7]This is a slight exaggeration. While Florida's Growth Management Act reads that way, in reality there may be ways to get a plan approved with unacceptable levels of service on roads. Dade County's plan was approved despite unacceptable levels of service under the justification that congested roads would help the county achieve the goal of using more mass transit. In this specific case, flexibility might be viewed as a virtue. In general, it means that the Department of Community Affairs has a large amount of discretionary authority over the local planning process.

[8]See Taub (1988) for a discussion of legal standing and the concurrency rule.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In summary, Florida's Growth Management Act will place all land in Florida under the jurisdiction of some local comprehensive plan. Development can take place only within the plan's parameters. Plans can be modified, but the concurrency requirement has the potential of placing a substantial roadblock in the way of any change, or of any development that would erode the level of service of infrastructure below the planned level, as defined in the comprehensive plan. The Act, therefore, has the effect of transferring some of the rights to use property from the property owner to the general public through the political process.[9]

Florida's Growth Management Act and Florida's Growth Problems

In general terms, the concepts of growth management and concurrency mean providing orderly growth and making sure that development does not outstrip the infrastructure that is servicing the development. The problem is difficult because infrastructure must be paid for as it is built, even though once in place, it will provide services for a long time. As written, Florida's Growth Management Act does not deal with this problem and actually makes one reasonable solution a violation of the law.

When stripped of the jargon about comprehensive plans, concurrency, and levels of service, Florida's Growth Management Act makes congestion of infrastructure illegal. Development cannot proceed without infrastructure in place or in progress to support it. But making congestion illegal does not specify how congestion is to be dealt with, and the Act does not provide any solutions. Florida's Growth Management Act deals with infrastructure the same way wage and price controls deal with inflation. Wage and price controls make inflation illegal but do not deal with the underlying causes of inflation. Likewise, the Act makes congestion illegal without dealing with the causes of congestion.

This observation is not new. Some individuals have commented that the Growth Management Act will force the legislature to deal with infrastructure problems, but this view may be optimistic. First, there are no easy solutions; otherwise the problems could have been dealt with directly in the Act rather than postponed. Second, until something is done to ease congestion, there is a conflict between the Growth Management Act and development to accommodate Florida's inevitable growth. These problems manifest themselves as polit-

[9]A good discussion of takings appears in Pilon (1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ical pressure on several fronts: pressure to reduce growth, pressure to increase funding for infrastructure, and pressure to modify the Act. Third, the wording of the Act prevents the financing of infrastructure out of the tax base that will be created by development.

One way to deal with congestion is to allow development that will increase congestion and lower levels of service. The development will generate more tax revenue, and the additional tax revenue can be used to produce more infrastructure. If infrastructure is created concurrent with development, then those who use the new infrastructure will be the occupants of the new development, but those new occupants have not been a part of the tax base. An alternative is to plan for future infrastructure by setting aside land for roads, parks, schools, and so forth, but then to wait to produce the infrastructure until the tax base is in place and paying taxes to finance the infrastructure. The result is a temporary increase in congestion until the infrastructure is completed.

This solution may not be appropriate in all cases, but the concurrency requirement in Florida's Growth Management Act makes it illegal in every case. A law that has good intentions does not necessarily produce good results. In this example, the Act outlaws one possible solution to the problems it is trying to solve.

Florida's Growth Management Act also encourages developers to develop overly rapidly to beat congestion problems. If development is illegal when the infrastructure is judged inadequate, then developers have an incentive to develop quickly before congestion occurs that will prevent future development. In this instance, the Act contains incentives to develop inefficiently.[10]

As communities grow, more population leads to more congestion. Residents tend to want the amenities of growth—the strong local economy produced by new jobs, good shopping areas, cultural activities, major league baseball—without the costs. In a growing community, residents may live in the same houses from year to year, but they are living in a changing community. One of the costs of living in a larger community is more congestion.

Property values rise as population in an area increases, and those higher property values reflect the increased locational desirability of property nearer the center of development. Existing property owners have an incentive to keep developed property scarce and to keep property values high by stifling development, but this behavior imposes a cost on the owners of undeveloped property who have

[10]For a discussion, see Wagner (1988) and Sonstelie and Portney (1978).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

some of the value of their property transferred—through artificially created scarcity—to owners of developed property.

Florida's Growth Management Act is against the best interest of renters (who tend to have lower incomes than homeowners) who will face higher rents, against the interest of future residents who will find higher housing costs whether they rent or buy, and against the interest of owners of undeveloped property who find it more costly to develop because of growth management regulations. These future residents are people who will migrate to Florida, but they are also the children of current Floridians who will want to purchase a home decades from now.

Florida's Growth Management Act purports to try to manage growth, but it sidesteps the most pressing problems of infrastructure and, as will be discussed below, encourages inefficient development patterns. In short, the Act does not effectively deal with Florida's growth problems, even though the language of the Act tries to eliminate some problems of growth and congestion just by making them illegal. The Act might be viewed as a statement of good intentions, but good intentions are not sufficient to manage growth. Some provisions of the Act will aggravate the state's growth problems.

Transfer of Rights through Florida's Growth Management Act

One of the potentially significant features of the concurrency requirement in Florida's Growth Management Act is the degree to which development can be challenged on concurrency grounds. From a legal standpoint, almost anyone has legal standing to oppose development on concurrency grounds.[11] The following Florida Statutes (F.S.) are relevant in this regard:

> (1) It is the intent of the legislature that substantially affected persons have the right to maintain administrative actions which assure that land development regulations implement and are consistent with the local comprehensive plan [F.S. 163.3213].
>
> (1) Any aggrieved or adversely affected party may maintain an action for injunctive or other relief against any local government to prevent such local government from taking any action on a development order, as defined in F.S. 163.3164, which materially alters the use of density or intensity of use on a particular piece of property that is not consistent with the comprehensive plan adopted under this part.

[11]See Taub (1988) on this point.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> (2) "Aggrieved or adversely affected party" means any person or local government which will suffer an adverse effect to an interest protected or furthered by the local government comprehensive plan, including interests related to health and safety, police and fire protection service systems, densities or intensities of development, transportation, facilities, health care facilities, equipment or services, or environmental or natural resources. The alleged adverse interest may be shared in common with other members of the community at large, but shall exceed in degree the general interest in community good shared by all persons [F.S. 163.3215].

The law clearly gives almost anyone the legal right to claim to be an "aggrieved or adversely affected party" and to stall any development to make sure that the development is in accordance with the local comprehensive plan. An individual who feels a new development will degrade levels of service on roadways, or will harm the environment, is explicitly given legal standing to oppose the development. Of course, the owners of developed property will oppose new development. Earlier discussion illustrates how owners of developed property can gain by opposing new development, and this discussion on the law shows that they clearly have the legal standing to oppose development on concurrency grounds.

Levels of Service

One of the key mechanisms by which Florida's Growth Management Act can manage growth is by applying levels of service as specified in the local comprehensive plans. The concurrency requirement applies to all infrastructure, but in practice the most significant and controversial aspects of the concurrency requirement deal with roads. One challenge in the planning process is defining the roadway capacity that is necessary to meet the concurrency requirement.

The local comprehensive plans specify levels of service to be provided by infrastructure, and local governments are able to set their own levels of service for infrastructure, providing that they receive state approval.[12] The levels of service specified in the comprehensive plans then define the levels of service required for concurrency. Any development that would reduce the level of service below that specified in the plan cannot be approved.

There has been extensive discussion about the merits of the Growth Management Act as a mechanism for producing orderly

[12]Brevard County was the first to submit a plan, but it was rejected by the state. For a discussion, see Winters (1989).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

growth.[13] The discussion seems to focus on building infrastructure to accompany development, but this strategy could miss the mark on two counts. First, additional roads may do little to lower congestion, and second, less-restrictive land use regulations could lower the burden on infrastructure.

The Amount of Infrastructure and Levels of Service

Roads are a scarce resource that are made more congested because, except in rare instances, there is no direct charge made for their use. Inefficient resource allocation results when scarce resources are owned in common and when anyone is allowed access at no cost. The inefficiency arises because nobody has an incentive to consider the costs their behavior imposes on others.[14] Drivers who enter congested highways impose costs on every other driver by slowing the progress of others, yet each individual driver has no incentive to consider the effects on others of entering the highway.

If all roads were toll roads, tolls could be charged to discourage use at peak hours. Ideally, the toll would rise during congested hours and would fall (perhaps to zero) when the road was not congested, thus discouraging use when the road is already crowded. In the absence of a toll, congestion acts as the only way to discourage travel at peak hours. Congestion is a rationing device, but it is not an efficient rationing device. Using congestion to ration roadway use gives preference to those who value their time the least, rather than those who value travel the most.

Congestion on roadways gives people an incentive to travel at off-peak hours. It also gives them an incentive to take fewer trips (doing more on each trip) and to live closer to where they work. Improving the roads will not necessarily reduce congestion, because better roads will entice more travel at peak hours, will encourage more driving, and will lower the cost of living far from one's workplace. Thus, there is not necessarily a direct relationship between the amount of infrastructure and the level of service provided.

Levels of service from a given infrastructure could be improved by encouraging development of work and shopping areas around the perimeter of urban areas rather than restricting it by zoning or other means to a central area. With a central work area, roads are congested in one direction in the morning as people go to work and in the other in the evening as they return home. Disbursed work areas create a

[13]See, for examples, 1,000 Friends of Florida (1988), Taub (1988), and Wilson (1989).
[14]An insightful discussion on efficient use of common resources appears in Cheung (1970).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

more even two-way flow of traffic and can improve the level of service from a given level of infrastructure. In an overly simplified model of commuting from suburbs to central city, population growth increases average commuting distance as suburbs are developed farther away, but if workplaces are spread around the perimeter of an urban area, people will find it easier to live near where they work, thus lessening commuting distances.

In Los Angeles, often considered to be one of the worst examples of urban sprawl, only 3 percent of the work force works downtown. Peter Gordon and Harry Richardson (1989) have identified 19 major activity centers in the Los Angeles area, but even these major activity areas account for only 17.5 percent of the area's jobs. The largest share of the Los Angeles work force lives and works in the suburbs, and those individuals face an average commute of 20 minutes. The burden on the infrastructure can be lessened by disbursing workplaces around the perimeter of an urban area rather than forcing everyone to commute to and from a central city, but ironically, zoning laws—and now comprehensive planning—will aid in producing additional congestion by preventing disbursed development.

Urban sprawl does not necessarily increase commuting distances; it may shorten them, if business development is not restricted from the perimeter of a city. Urban sprawl also might lessen the environmental impact of development if it allows more of the natural environment to remain intact. High-density housing will typically eliminate the natural landscape, for example, while a house on a two-acre lot will have little impact on the environment. In some areas of Florida, stormwater runoff is a problem, but while almost all water becomes runoff in high-density areas, low-density development can be designed to produce almost no runoff and, in some cases, might even absorb runoff from high-density areas. In this case, urban sprawl has the potential to help solve one of problems of development. One of the goals of comprehensive planning is to "discourage the proliferation of urban sprawl,"[15] but as this paragraph suggests, urban sprawl is not unequivocally undesirable when compared to the alternative of high-density urban areas.

The Governor's Task Force on Urban Growth Patterns clearly considers urban sprawl to be a problem. According to the Task Force, "most of Florida's future growth will be accommodated through sprawling, low-density development on raw land, with jobs and housing moving ever away from existing urban centers, unless decisive action is taken at every level of government and by the private

[15]Florida Administrative Code Rule 9J.5.006(b)7.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sector to reverse this trend and promote efficient, compact urban development patterns" (Rotella et al. 1989, p. 1). This urban sprawl can reduce commuting distances, can lessen environmental impacts, and can afford better lifestyles to Florida's residents if workplaces can develop outside urban centers. The report illustrates that the population density of Florida cities has been declining over several decades, but normally, having the population live in less-crowded conditions is considered an improvement in the standard of living. Unquestionably, there are drawbacks to having more urban centers of smaller size and to having lower-density development. But while recognizing these costs, many critics of urban sprawl overlook the potential benefits.

The State of Florida has contributed to urban sprawl around Florida's major cities. The University of South Florida in Tampa and the University of Central Florida in Orlando were both built a substantial distance away from population centers in those cities, yet development after the establishment of those universities has allowed those in the university communities to live closer to the universities, reducing commuting and the associated burden on the infrastructure. In contrast, Florida State University is in downtown Tallahassee, causing university commuters to compete with other downtown workers for space on the roads, adding to congestion. Those who oppose urban sprawl ought to argue that the location of FSU in Tallahassee is more desirable than that of USF and UCF in Tampa and Orlando. Would the residents of Tampa and Orlando really rather have those universities located downtown to reduce urban sprawl?

People who like things the way they are today would view urban sprawl as undesirable, but in a growing state the alternative to urban sprawl is not the way things are today. The question is whether future growth will shoehorn more people into urban areas and increase population density or will allow disbursed growth with less crowding. Critics of urban sprawl who live in relatively large houses on big lots and who might be reluctant to move into more crowded conditions themselves should consider why they believe it desirable for others to live in those conditions.

Lower-density development may or may not be desirable in specific cases, but public policy in Florida has taken the side against lower-density development without recognizing that it can bring some benefits. If this policy is carried out, it will prevent more disbursed development when such would be desirable and will add to the congestion problem. Public policy ought to weigh the costs against the benefits when considering development density, rather than being automatically biased against low-density development. It

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

is clear that a policy that unequivocally opposes urban sprawl is not in the best interest of either the citizens of Florida or the environment in general.

This section illustrates some of the conceptual problems with, and some alternatives to, attacking the concurrency problem by using tax dollars to build infrastructure. There is not necessarily a direct relationship between money spent on road construction and the amount of congestion on roadways, and policies designed to solve infrastructure inadequacies simply by spending more money will not necessarily succeed. If this point is obvious, it has not been made clear in some of the debate on the issue, which centers on where the money will come from to finance improvements in infrastructure.[16]

Goals versus Results: A Preliminary Assessment

Florida's Growth Management Act is only a few years old, so it is too early to assess the performance of the Act. It is, however, possible to make a preliminary assessment on the basis of the stated goals of the Act and of the likely results of the Act's provisions that were intended to further those goals. Many, but not all, of the stated goals of growth management are uncontroversial. But the Act is structured in a way that is likely to lead to unintended negative consequences.

Most people would agree with the broad goals of producing orderly growth, of protecting the environment, and of producing infrastructure concurrent with development. The goal of preventing urban sprawl is more controversial, at least partly because urban sprawl has not been precisely defined.[17] Given that the state is growing, it is not clear that in all cases high-density development is more desirable than low-density development; there are strong arguments to be made for the benefits of smaller high-density areas disbursed throughout an area as opposed to a single, concentrated, high-density area surrounded by suburbs. Urban sprawl has a negative connotation, but for a given amount of growth, lower-density development is not necessarily less desirable than higher-density development.

Another important issue is the legitimate question about the way in which infrastructure demands ought to be met. Simply building more roads in already congested areas might result in little visible

[16]Taub (1988), for instance, contemplates whether the concurrency doctrine will be an effective moratorium on development or will provide the impetus to fund infrastructure improvements.

[17]See, however, the Florida Department of Community Affairs *Technical Memo*, Vol. 4, no. 4 (undated), which does attempt to identify both the undesirable characteristics of urban sprawl and what local governments can do to avoid them in local comprehensive plans.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

impact because, without charging for access to those roads, additional capacity will simply encourage more use from those already in the area. Congestion acts as a rationing device; adding capacity encourages more use at peak hours, encourages more frequent travel, and reduces the incentive to live close to where one works. Yet the Growth Management Act's emphasis on concurrency has focused the discussion on where the resources will come from to fund infrastructure rather than on innovative approaches to solving the infrastructure problem.

When resources are allocated through the private sector of the economy, distributional issues are secondary because all parties have an incentive to structure trades so that everyone gains. Otherwise, trade would not take place. When resources are allocated through the public sector, distributional issues are important because, unless unanimous consent is required, there is always the possibility that some individuals will impose costs on other nonconsenting individuals through the political process.[18] In the democratic political process, property rights are poorly defined. As such, individuals can use the political process to benefit themselves at the expense of others. This is why the distributional aspects of growth management are likely to be the driving force behind any actual policies that are implemented.

While the stated intentions of Florida's Growth Management Act are in large part desirable, the consequences of the Act will not measure up to its goals. If the Act is enforced as it is written, it will transfer private property rights from property owners to those who are politically active, it will slow growth, increase the value of developed property, and decrease the value of undeveloped property. It will make property inside designated urban areas more valuable, and property outside areas designated for development less valuable. It will create political conflict because any development can be challenged within the concurrency requirement of the Act.

If Florida's Growth Management Act is strictly interpreted, rents and housing prices in Florida will rise as a result. Unfortunately, the connection between growth management and housing prices is not likely to be clearly drawn in political debate. Thus, proponents of affordable housing are not likely to attribute skyrocketing housing prices to the statewide regulation of land use mandated in the Act. In situations like this, the result is often additional political action,

[18]Buchanan (1976) discusses the fiscal exchange model of taxation, which posits that taxes are a price paid for government output. This model implies that those using the government output should be the same ones who pay for it.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

such as subsidized housing or rent controls, to offset the unintended negative results of the initial political action.

If Florida's Growth Management Act is loosely interpreted, the negative consequences could be limited to the costs of drawing up local comprehensive plans, the modification of those plans, and the legal expenses of developers and of the state in complying with the law. This outcome would be desirable compared to the economic and political costs that would result from a strict interpretation of the Act.

Conclusion

While this paper deals specifically with Florida's experience with growth management, there are broader lessons that apply to growth management movements in any state. Fundamentally, what growth management means is taking some rights away from the nominal owners of property and making them subject to the political decision-making process. As a result, rights that at one time were clearly defined become more poorly defined, and the problem of common ownership arises.

If the owner of a piece of property has the right to develop it, then the market provides incentives for the property to be developed in its most highly valued use. When the development of property is contingent upon political approval, the same market signals are not as effective. This reasoning applies to statewide growth management, as exists in Florida and Oregon, and also to zoning and other restrictions on the use of property. With growth management in general, and with Florida's concurrency rule in particular, poorly defined rights to develop property produce an incentive to develop too rapidly. If the possibility of developing a piece of property exists at the present but may be taken away in the future, then the developer has an incentive to develop now rather than risk losing the right. Ironically, laws designed to control development can perversely encourage development that is inefficiently rapid. The problem arises from a transfer of private property rights into common ownership through growth management legislation.

The economic justification for growth management legislation is that externalities produced by development should be controlled by government, but because of the common ownership problem, growth management legislation is a poor method of control. Private property owners have the incentive to use their property in its most highly valued use, and public planning is unlikely to improve upon these private decisions (Siegan 1970, 1972). If it is in the public interest to

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

leave some environmentally sensitive land undeveloped, or if some land should be used for parks or other public purposes, then the solution is for the government to buy it, rather than regulate its use by private owners. But note that if, for example, parks are socially desirable in neighborhoods, developers themselves have the incentive to produce them in order to increase the market value of the other land they develop and sell.

By taking the right to develop away from the nominal property owner and making it a political decision, inefficient development patterns are more likely because market signals are absent or distorted. As argued above, the standard plan for development in Florida will produce suboptimal land use patterns and will result in more traffic congestion than if the market were allowed to dictate land use patterns. In an effort to prevent "urban sprawl," Florida planners will force more distant commutes from suburbs to central cities by preventing business and commercial development outside already existing central urban areas. In general, land use planning is a method for overriding market signals about the efficient use of land and thus, in general, will produce inefficient development.

Fundamentally, growth management laws are a method of transferring some ownership rights of property from some people to others. One way that a person can acquire the right to determine the use of a piece of property is to buy it. But another way is to pass a law that transfers that right into the political process and then uses political power to dictate the way the property is used.[19] Growth management is popular because it allows the proponents of the legislation to dictate how land is used by taking the right away from its current owners rather than buying it. But it is inefficient because the resulting right is owned in common rather than privately owned. Economists are well aware of the inefficiencies that arise from common ownership, such as is produced by growth management legislation.

While this paper has focused on growth management legislation as it has affected Florida, similar laws exist in other states, and there is the potential for such legislation everywhere. Perhaps a case study of Florida's Growth Management Act can provide a general warning. Legislation intended to get people to use their property in a publicly responsible way rather than for purely private gain has a nice ring to it. With growth management legislation, this means appropriating peoples' property and overriding the market in a way that is both unfair and inefficient. The results of the law are quite different from their stated intentions.

[19]The result is taxation by regulation, to use the words of Posner (1971).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

References

Advisory Commission on Intergovernmental Relations. *Residential Community Associations: Private Governments in the Intergovernmental System?* Washington, D.C.: ACIR, May 1989.

Alchian, Armen. "Some Economics of Property Rights." *Il Politico* 30, no. 4 (1965): 816–29.

Boudreaux, Donald J., and Holcombe, Randall G. "Government by Contract." *Public Finance Quarterly* 17, no. 3 (July 1989): 264–80.

Boudreaux, Donald J., and Holcombe, Randall G. "Contractual Governments in Theory and Practice." In *The Voluntary City*. Edited by David Beito. Forthcoming.

Buchanan, James M. "Taxation in Fiscal Exchange." *Journal of Public Economics* 6 (July/August 1976): 17–29.

Cheung, Steven N. S. "The Structure of a Contract and the Theory of a Non-Exclusive Resource." *Journal of Law and Economics* 13 (April 1970): 49–70.

Florida Department of Community Affairs, *Technical Memo*, Vol. 4, no. 4 (undated, released in 1989).

Gordon, Peter, and Richardson, Harry. "You Can't Get There from Here." *Reason* (August/September 1989): 34–37.

Holcombe, Randall G. "Distributional Aspects of Florida's Concurrency Requirement." *Florida Policy Review* 5 (Winter 1990): 8–14.

1,000 Friends of Florida. "Concurrency Key to Growth Management." *Foresight* 1 (December 1988): 1.

Pilon, Roger. "Property Rights, Takings, and a Free Society." In *Public Choice and Constitutional Economics*, chap. 8. Edited by James D. Gwartney and Richard E. Wagner. Greenwich, Conn.: JAI Press, 1988.

Posner, Richard A. "Taxation by Regulation." *Bell Journal of Economics and Management Science* 2 (Spring 1971): 22–50.

Rotella, Ron, et al. "Governor's Task Force on Urban Growth Patterns, Interim Report." Tallahassee, Fla., March 1989.

Siegan, Bernard H. "Nonzoning in Houston." *Journal of Law and Economics* 13, no. 1 (April 1970): 71–147.

Siegan, Bernard H. *Land Use Without Zoning*. Lexington, Mass.: D.C. Heath & Co., 1972.

Sonstelie, Jon C., and Portney, Paul R. "Profit Maximizing Communities and the Theory of Local Public Expenditure." *Journal of Urban Economics* 5 (April 1978): 263–77.

Taub, Theodore C. "Florida's Growth Management Concurrency Doctrine—Moratorium or Impetus to Fund Needed Infrastructure." *Environmental and Urban Issues* 15 (Fall 1988): 5–12.

Wagner, Richard E. "Agency, Economic Calculation, and Constitutional Construction." In *The Political Economy of Rent-Seeking*, chap. 29. Edited by Charles K. Rowley, Robert D. Tollison, and Gordon Tullock. Boston: Kluwer, 1988.

Wilson, Elisabeth. "The Coming Backlash." *Florida Trend* (January 1989): 54–58.

Winters, Jan. "Approaches to Growth Management in Palm Beach County." *Environmental and Urban Issues* 16 (Winter 1989): 7–12.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Land Use Regulations Should Preserve Only Vital and Pressing Governmental Interests

Bernard H. Siegan

The Failure of Planning

Land is a precious and scarce natural resource. It should be used to best provide for the needs and desires of the people. This objective will best be realized if the use and development of land is left to the private marketplace, except in those instances when government has a vital and pressing need to impose regulation. The great lesson of our times is that the forces of production, conservation, and creativity rest principally in the marketplace and not in government. True, private entrepreneurs act largely in their own self-interest, but probably no more so than people in government, and their endeavors in the economic area are much more oriented toward the general welfare.[1]

This wisdom is now subscribed to by government leaders throughout the world including those in China, the Soviet Union, and other Eastern bloc countries.[2] Zbigniew Brzezinski, who served during the

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Distinguished Professor of Law at the University of San Diego. This article is an edited version of the keynote address delivered by the author on February 28, 1989, at the Fifth International Congress of Urbanism held in Porto Alegre, Brazil. A similar article will appear in the *University of San Diego Law Review*.

[1]This view is consistent with the "public choice perspective" advanced by James M. Buchanan, Nobel Laureate in Economics for 1986. He emphasizes self-interest as the motivating factor in both private and political choice. However, the forces of the economics marketplace are more likely to channel individual self-interest into socially desireable outcomes. See, for example, Buchanan and Tullock (1962), and Buchanan (1987).

[2]Indicative of the greater acceptance of capitalist theory, 50 countries during the past 10 years—including most major industrial countries—have significantly reduced their maximum marginal tax rates on individual incomes. In the United States, this rate was 70 percent in 1979 and 28 percent in 1989 (Reynolds 1989).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Carter administration as assistant to the president, National Security Affairs, and as director of the National Security Council, recently wrote: "Prospering only where it abandons its internal substance even if still retaining some of its external labels, communism will be remembered largely as the twentieth century's most extraordinary political and intellectual aberration" (Brzezinski 1989, p. 1).

Leaders in the Marxist world have come to recognize that private ownership and enterprise offer their countries relief from economic regression and stagnation. These leaders understand that collectivist systems have wasted and exploited their nations' resources, and that their only hope for economic and social progress rests with private enterprise and ownership.[3]

Such philosophical conversions are understandable. By comparison to the communist world, capitalism's economic accomplishments have been enormous. Some perceptive observers explain that communism has eliminated the booms and busts of capitalism by eliminating the booms. According to Brzezinski (1989, p. 67), in recent years, research and academic economists supporting Gorbachev urged a "truly comprehensive dismantling of the centralized planning apparatus. . . . Central planning, they argued, had to be replaced with a modified market mechanism."[4]

Leo Aron (1989, pp. 30–31), who formerly lived in the Soviet Union, writes that, since Glasnost, reports within the Soviet Union have been candid about its economic problems. With the official poverty level set at 75 rubles per person per month, it has been reported that 43 million people are below the poverty line, and 40 percent of Soviet families, about 100 million people, live on less than 100 rubles a month. The consumption of meat and dairy products by the Soviet poor has declined by 30 percent since 1970. A total of 1.2 million beds are in hospitals with no hot water, every sixth bed is in a hospital with no running water at all, and 30 percent of Soviet hospitals do not have indoor toilets. Half the nation's elementary schools lack central heating, running water, or sewage systems.

According to Brzezinski (1989, p. 238), in 1983 there was one car for every 1.8 Americans, 4.4 Japanese, 2.5 Germans, and 2.8 Italians, but only one for every 14.2 Soviets, 5.8 Czechs, and 10.8 Poles. He stated, "It is a startling fact that blacks own more cars per capita in

[3]See generally Brzezinski (1989).

[4]See generally Brzezinski (1989) and Shelton (1989). For some interesting insights into the operation of the Soviet economic system, see Soviet leader Mikhail Gorbachev's recent book (1987, pp. 251–302). Specifically, note his criticisms of central planning (pp. 279–85) and his conclusion (p. 284) that a "serious restructuring is required in planning work."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

South Africa than do citizens in the Soviet Union." In October 1989, the *San Diego Union* (1989, p. A2) reported that Soviet consumers were suffering some of the worst shortages of food and other goods the country had ever known. "With the approach of winter, complaints ha[d] grown to an angry chorus."

The introduction of capitalist programs has had remarkable results. Initiated in 1978, the progressive decollectivization of Chinese agricultural land "had prompted a dramatic surge in productivity. . . . China was transformed from a net importer of food to an actual exporter" (Brzezinski 1989, p. 165). "This agricultural rebirth stimulated growth in the output of Chinese rural industry, which increased by a staggering 400 percent between 1981 and 1986 and which grew by a further 36 percent in 1987 alone" (Brzezinski 1989, p. 166).

Capitalism's accomplishments are well illustrated by the state of the U.S. economy in recent years. Notwithstanding the minimal amount of price controls in the United States, prices rose annually during the seven years including and prior to 1988 an average of about 3.6 percent, the biggest increase occurring in 1988 in the amount of 4.4 percent and the lowest in 1986 at 1.1 percent.[5] Moreover, during this period the economy was strong, registering a record length of prosperity. It grew at an average annual rate of about 3 percent, and unemployment decreased from 9.5 percent to an historically low rate of 5.4 percent.[6] This impressive record is in large measure attributable to the high level of production that existed during the period.

To be sure, communism's failures do not mean that every control over private enterprise or property will be unsuccessful. However, the magnitude and breadth of that system's failures should encourage careful inquiry into the costs and benefits of limitations on private economic activity. At the very least, it would seem that the United States should be wary of using economic restrictions that have not worked in the communist countries. Thus, one major failing of the communist system has been master planning. On a micro level, mas-

[5]See Council of Economic Advisers (1989, p. 24). The foregoing figures do not take into account the improvement in quality that occurs over time. Were quality improvement considered, the price increases would be, in my opinion, significantly less than those set forth. Thus, it has been suggested that a 2 percent per year increase in the official price index would be the equivalent of stable prices ("An Interview with Milton Friedman" 1989, p. 7).

[6]See Economic Report of the President (1990, pp. 294, 301) and Council of Economic Advisers (1989, p. 12). For comparison purposes, the author selected the said 7-year period of this nation's economic history because the Reagan administration (then in power) sought to achieve the capitalist objectives of less taxation, less regulation, and less spending.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ter planning is an essential part of zoning regulation. Under zoning, the use of all property in a locality is restricted pursuant to a master plan. The major trouble in zoning—as elsewhere—is that the plan limits production arbitrarily with adverse effects on supply, prices, and competition.

The President's Commission on Housing

Eliminating master planning is a theme common to the recommendations for reform of land use regulation as proposed by President Reagan's Commission on Housing. The President appointed the 30-member commission in 1981 to seek ways to increase the supply of housing and lower its cost. The commission's report[7] issued in 1982 urged severe limitations on zoning and other land use regulations in order to accomplish these purposes. It urged that municipalities retain only such powers over the development of private housing as are necessary to preserve vital and pressing government interests.

The report contains more than 100 recommendations covering all aspects of housing. The proposals on land use present a message common to many others endorsed in the report: To obtain more housing at lower cost, remove most government controls. The commission recommended that a "vital and pressing" standard be applied to determine the need for regulation. According to the report, market forces should determine the location and composition of new development—type, size of unit, density, land coverage, and height—unless the municipality could prove that it had, in effect, an extremely persuasive justification—a vital and pressing one—for applying restraint. The objective is to forbid imposition of regulations that curtail housing development and that do not, in fact, implement police powers.

Similar to the determinations made by presidential commissions appointed by prior administrations,[8] Reagan's Commission (*Report* 1982, pp. 177–82) found great wrongs and abuses at the local level but, unlike the prior studies, proposed deregulation. Regulation, this commission concluded, was the problem, not the solution. Deregula-

[7]The recommendations on regulations are contained in Section IV of the *Report of the President's Commission on Housing* (1982) [hereinafter called *Report* 1982], beginning on page 177. The author of this article was a member of the commission, serving as chairman of the Regulations Committee.

[8]See the President's Commission on Urban Housing, *A Decent Home* (1969) and the Report of National Commission on Urban Problems to the Congress and to the President, *Building the American City* (1968, p. 20). The latter states, "In short, although the basic justification for zoning is to protect the overall public good, this often appears to be the last consideration as zoning is now practiced."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tion not only would remove the wrongs and abuses, but also would better serve the public interest.

Some local controls relating to use and development should continue. Only municipalities are able to plan and build streets, parks, public buildings, schools, storm and sanitary sewers, and water mains. Municipalities must also secure the public's vital and pressing interests, which the commission (*Report* 1982, p. 200) identified broadly as limited to protecting health and safety, remedying unique environmental problems, preserving historical resources, and protecting investments in existing public infrastructure resources. More specifically, "vital and pressing governmental interests" include requiring adequate sanitary sewer and water services and flood protection, ensuring that topographic conditions will permit safe construction and accommodate septic tank effluents, protecting drinking water aquifers, avoiding nuisance or obnoxious uses, requiring off-street parking, prohibiting residential construction amid industrial development, and avoiding long-term damage to the vitality of historically established neighborhoods.

The commission's recommendations (*Report* 1982) apply only to the development of housing. The vital and pressing standard is set forth as follows:

> To protect property rights and to increase the production of housing and lower its cost, all State and local legislatures should enact legislation providing that no zoning regulations denying or limiting the development of housing should be deemed valid unless their existence or adoption is necessary to achieve a vital and pressing governmental interest. In litigation, the governmental body seeking to maintain or impose the regulation should bear the burden for proving it complies with the foregoing standard.[9]

Accordingly, government controls would be limited to forbidding private activity that is truly harmful, much as is the case when other private actions such as speech, press, religion, mobility, or privacy are restricted.[10] The proper role of government is to protect the public from harm and not to prohibit or diminish production, competition, and creativity.

Protecting Private Rights

Residents often complain when new development increases traffic and pollution of air and water, and when it decreases open space.

[9]The commission adopted this recommendation without any dissent.

[10]The commission suggested an elevation in the level of judicial scrutiny applied in land-use decisions to provide greater judicial protection for rights of ownership (*Report* 1982, p. 202 n.13).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

They object that growth brings more crime and taxes. Although comfort and convenience levels may be reduced for some as a result of growth, it is difficult to conclude that existing residents have been injured in a legal sense. In our society, states and municipalities are part of a political union and not walled enclosures serving solely their own interests. They are expected to develop and grow to serve the diverse demands of the population. To exclude people and limit development to a level that is not consistent with these ordinary expectations would unfairly grant special benefits to certain residents. As a New York judge observed many years ago in *Bove v. Donner-Hanna*,[11] one who chooses

> to live in the large centers of population cannot expect the quiet of the country. Congested centers are seldom free from smoke, odors, and other pollution from houses, shops, and factories, and one who moves into such a region cannot hope to find the pure air of the village or outlying district. A person who prefers the advantages of community life must expect to experience some of the resulting inconveniences.

In *Edwards v. California* (314 U.S. 160 [1941]), the U.S. Supreme Court struck down a California statute that made criminal the bringing of nonresident indigents into the state. In support of this law, California argued that the influx of migrants had resulted in staggering problems of health, morals, and especially finance. However, the Court (at 173) held that the law violated the right of travel protected under the commerce clause of the Constitution; no "state [may] isolate itself from the difficulties common to all of them by restraining the transportation of persons and property across borders." The Court (at 174) noted that indigent nonresidents were deprived of the opportunity to exert political pressure on California lawmakers in order to obtain a change in policy. The same also occurs in growth control situations.[12]

The Pennsylvania Supreme Court in *National Land and Investment Co. v. Easttown Township Board of Adjustment*[13] viewed the matter of limiting development as one of individual liberty:

[11]Bove v. Donner-Hanna Coke Corporation, 236 A.D. 37, 40, 58 N.Y.S. 229, 232 (1932).

[12]It has been held that state laws that deny newly arrived indigents welfare benefits during their first year in the state inhibit the indigents' exercise of the right to travel. Thus, such laws are unconstitutional according to Shapiro v. Thompson (394 U.S. 618 [1969]). Growth controls operate similarly and may fail similar constitutional examination.

[13]419 Pa. 504, 532, 215 A.2d 597, 612 (1965). The right involved in National Land v. Easttown has been identified by one federal judge as the right to travel. See Construction Indus. Ass'n v. City of Petaluma, 375 F.Supp. 574 (N.D. Cal. 1974), *rev'd on other grounds*, 522 F.2d 897 (9th Cir. 1975), *cert. denied*, 424 U.S. 934 (1976).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> The question posed is whether the township can stand in the way of the natural forces which send our growing population into hitherto undeveloped areas in search of a comfortable place to live. We have concluded not. A zoning ordinance whose primary purpose is to prevent the entrance of newcomers in order to avoid future burdens, economic and otherwise, upon the administration of public services and facilities cannot be held valid.

Erecting barriers to travel and occupancy interferes with a free society's ideals of mobility and opportunity. Without a vital and pressing justification, growth control lacks equitable or philosophical roots in a legal system that is essentially dedicated to maximizing liberty.

Actually, as previously stated, the vital and pressing standard does allow a community to limit development in order to remedy "unique environmental problems." The burden would be on the community to prove that conditions have reached an unusually adverse level demanding application of growth restrictions. In addition, the community must show that the primary purpose of the regulation is to ameliorate this problem and not to exclude people since, as the Reagan Commission (*Report* 1982, p. 200) concluded, "Exclusion is clearly not an acceptable governmental interest."

Costs and Benefits of Development

Some people reject growth on a utilitarian basis: It does not pay its way; therefore, growth raises taxes. Many studies have been conducted to determine the costs and benefits of new development, and the results are mixed. The problem with such studies is they are necessarily incomplete and misleading—they are unable to quantify many advantages that localities obtain from new residents. For example, how does one evaluate the gain to the elderly from the arrival of a cancer or heart specialist or nurses to staff the local hospital, all of whom might be excluded by growth controls or by the increased cost of housing such controls cause? Similar advantages are reaped from other professionals or technicians, who might also be excluded. Larger cities boast of cultural, educational, religious, library, medical, shopping, and transportation facilities not available in smaller cities. Growth controls can also be costly when they cause unemployment or reduce the incomes of those who depend for their livelihoods directly or indirectly on the construction industry, one of the most important industries in the nation.

Some people demand that in addition to paying for directly related costs, new developments should also pay for all indirect costs, such as the added use of roads, highways, sewer and water facilities,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

schools, libraries, and so on. However, these facilities are the property of the community and all its residents, both new and old. Congested roads may become more congested with the arrival of newcomers, but the problem is not solely attributable to them: It results from development of the entire community. If the community had not developed so extensively, the newcomers' impact would not be significant.

Once a person becomes a resident of the city, he or she is entitled to all benefits and burdens of residency—the use of facilities as well as the responsibility of paying for those facilities. If their homes are specially assessed (either before or after they purchase them), new residents would be paying, in effect, an entry fee, which they would again have to pay in taxes as residents. They would be subject to double taxation.

Deregulation of zoning may augment tax receipts. By limiting zoning restrictions, cities and towns will encourage the location of developments that are highly profitable for the tax base, such as factories, shopping centers, and high-rises, all of which yield substantial tax revenues and cost relatively little to service. The plants and centers result in no children and the high-rises, relatively few children; schools are by far the most expensive of municipal services, typically accounting for two-thirds to three-fourths of real estate tax bills.

By limiting growth, towns and cities also limit the number of potential shoppers, consequently deterring the construction of shopping centers. Housing limitations may likewise exclude workers for industry and thereby prevent location of that revenue source in the area.

Planning and Politics

Areawide land use controls arrived in this nation in 1916 in the form of the New York zoning resolution, the country's first zoning ordinance. This modest ordinance contained three use districts: residential, commercial, and unrestricted; five classes of height districts; and three classes of area districts.[14] New York now has about 70 zoning districts and a host of other controls never conceived of by the framers of its original ordinance. A similar story applies to other cities.

Small, modest ordinances grow into very complex and complicated ones. One reason is, of course, the change in conditions, techniques,

[14]See Toll (1969, pp. 172–87).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and thinking that occurs over the years and that is reflected in our laws. But there are two other explanations for the uncontrolled growth of zoning. The first is that zoning has been the story of unrealized expectations. To date, we have had six or seven different zoning systems or strategies in this country. Each has been introduced with what has turned out to be greatly inflated rhetoric as to what it would accomplish. Toll (1969, p. 181) describes the reception that New York's 1916 zoning resolution obtained:

> The most unusual thing about the new law was the passionate optimism with which it was received, not only locally, but throughout the nation. This spirit swept far beyond the young planning profession and the business community. McAneny [President of the Board of Aldermen] called the law "the greatest single achievement in city planning in America," and "the greatest thing New York City . . . has ever done. . . ." For *World's Work* it opened "a new era in urban civilization." *Outlook* hailed it as "one of the most progressive and forward-looking steps for the protection of its future development that has been made by any American city." Veiller, the hard-bitten, rather cynical reformer, said: "The situation as I see it is the most hopeful one the city planning group has ever had to face. I repeat, we are standing on the edge of a great change in living conditions in America. We are going to revolutionize conditions in a generation. . . ."

Each zoning system, in turn, has for the most part failed to meet the expectations created by that rhetoric. The result, every time, is a new effort at the drawing boards, producing more and increasingly severe rules and regulations that, experience suggests, are not likely to be more successful than the previous ones.

Another reason for the proliferation in zoning regulations is that the process is basically one of resolving differences among various special interest groups in the community. No matter how perfect the zoning plan, it will help some people and hurt others. Soon after passage of the ordinance, the losers, experience shows, start doing those things that will make them winners. Owners will seek to rezone their property to increase its value. Homeowners and conservationists will lobby to "downzone" property, and civic groups will move to make their reforms. The courts may also effect significant changes. Before long the original plan will be largely reduced to history because of innumerable amendments grafted to it.

The dominant factors in zoning have been public pressures and political influences—the factors to which officeholders who control zoning will respond. Matters such as efficiency, demand, and environment have been subordinate and often nonexistent considerations unless they coincide with political pressures.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Public planning of land use is highly subjective and lacks those standards and measurements that are requisites of a scientific discipline. Planners confront serious problems in evaluating the present and forecasting the future, whether it be on a micro or macro level. With respect to zoning, overruling market-based decisions on land use would seem to require an adherence to special goals or values, or perhaps an understanding that relatively few possess. Zoning experience is replete with instances where planners have classified land either to allow uses unacceptable in the market or to deny uses eagerly sought in the market. Theory and education alone cannot substitute for the actual experience of making practical decisions and suffering their consequences. Few planners have ever been part of the construction or development industry; they have not been responsible for decisions in locating and developing residential, commercial, or industrial projects. Even if they once had been, their information about prices, materials, innovations, trends, and consumer desires and preferences must necessarily now come from secondary or more remote sources, not directly from the firing line.

How then can planners possibly be as familiar with the location, development, construction, and operation of shopping centers, housing developments, nursing homes, or mobile parks as those who develop, own, and operate them? Owners and their mortgage lenders risk substantial funds to achieve success. Master planning inevitably will forbid developers to build where they want to and will permit them to build where they do not want to.

However, regardless of their knowledge or ability, the fact is that planners are not destined to make a significant impact on the regulation of land use. Major decisions will be adopted by elected officeholders who possess the final power. They can be expected to, and will, respond to a variety of pressures and concerns, a principal one being the interests of people who help them obtain and keep public office.

The question arises: Can the widely recognized problems of land-use regulation be solved by elevating the level of control from local to state and federal authorities? The simple answer is that political pressures, moral and legal corruption, and bad laws are not confined to local government. A second answer comes from the experience of regulation in this country. That record has been an extremely poor one. Of the numerous studies made of regulatory agencies, many have been printed in the *Journal of Law and Economics*, published by the University of Chicago Law School. A former editor, Ronald Coase (1975, p. 169), concludes as follows:

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> The main lesson to be drawn from these studies is clear; they all tend to suggest that the regulation is either ineffective or when it has a noticeable impact, that on balance the effect is bad, so that consumers obtain a worse product or a higher priced product or both, as a result of the legislation. Indeed, this result is found so uniformly as to create a puzzle; one would expect to find in all these studies at least some government programs that do more good than harm.

Accordingly, the history of zoning and other economic regulation in this country provides very strong support for limiting regulation of land use pursuant to the vital and pressing standard previously discussed.

State Environmental Regulation

In recent years, considerable attention has focused on imposing state regulation to prevent development in areas considered environmentally sensitive. One major problem is defining environmentally sensitive land. Thus, land that is miles from the California coast was included in the territory controlled under that state's coastal initiative. However, regardless of the language or intent of its sponsors, it is not likely that such legislation will or can prohibit development within the target areas. In all probability, what will happen is much more of what is already occurring under local zoning: less development, more land for urban purposes, and increases in real estate prices and rents. Relatively few critical areas would be preserved in their natural setting.

There are two basic reasons for this. First, there are the "taking" and due process provisions of federal and state constitutions that limit the extent to which private property can be regulated. This paper's next section will discuss recent U.S. Supreme Court decisions in this area. Second, the regulatory process tends to resolve controversies through some compromise formula, an approach that would be reinforced in this situation by the said constitutional provisions.

The history of California's coastline regulation, established in 1972 by public initiative, discloses how land use controls would probably operate in areas designated as environmentally sensitive. In 1973, more than 6,200 permit applications were received by the six regional commissions given permit power under this law. Of those applications, about 5,200, or 83 percent, were granted in whole or in part. In 1974, of the approximately 4,700 permit applications received, some 4,400, or 94 percent, were approved in whole or in part. Such figures do not fully reflect the regulation's impact, however.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

To obtain a better understanding of what these numbers mean, consider Bruce Johnson's analysis (1974, p. 18) of the operation of the Santa Barbara Coastal Commission, on which he served. During the first 14 months of that commission, it approved 95 percent of the applications received for single-family dwellings, but it allowed only 60 percent of the single-family units requested on these applications. While 77 percent of the multifamily applications received approval, only 51 percent of the units applied for were accepted.

Accordingly, we may conclude that if a developer applies for a permit to built 100 units on his five-acre tract adjoining the Pacific Ocean, he will, after considerable delay, red tape, political maneuvering, and possibly a nervous breakdown, have his request whittled down to say, around 60 units. That is what is referred to in regulation as "balancing."

The five acres in our example will still remain in private ownership and will not normally be accessible to the public, except possibly for an easement to the coastline that would be required, but would have more theoretical than practical value. It is questionable that better views would be provided since the buildings might also have to be aesthetically compromised. The question that now presents itself is what happens to the 40 units that were disallowed? One possible answer is that those, along with other disallowed units, represent unsatisfied demand and will, in time, cause greater or more rapid development of the coast, something entirely inconsistent with the aims of the initiative. It is, of course, possible that the 40 units will never be built, in which case society will have lost the benefits of considerable private expenditures benefiting business, employment, tax revenues, and housing supply, and there will be less housing at higher cost.

Possibly those 40 units may find their way into existing vacant land within a developed city. But they will reduce the land available for other city projects. Another alternative is that the disallowed units will increase demand for building in rural, undeveloped areas at the expense either of other environmentally sensitive land, or of land used or suitable for agriculture, grazing, and mining. In this case the result will be more spread and sprawl, exactly those horrors from which regulation is supposed to protect us. But this consequence is not an unusual one these days. Slow- and controlled-growth policies have caused development to leapfrog the areas so restricted and to sprawl and spread farther and farther into areas that are or should be used for other purposes.

An article in the *Texas Monthly* (Burka 1977, p. 109) describes the situation very succinctly. It concludes the following about the growth policies of the Austin (Texas) City Council:

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> The irony is that the most anti-growth council in Austin history may have done more to bring about urban sprawl than did any of the pro-developer councils that preceded it. Over 75 percent of the homes built in the Austin metropolitan area last year were outside the city's extraterritorial jurisdiction.

The relation of these experiences to other proposed statewide land use controls should be apparent, and much more than just the California coastline would be involved. Development would take place, but with less intensity of use; while there would be more open space, it would be privately owned and not normally accessible to the public. The proposed state controls would accomplish more of what the proponents say they want to prevent; instead of conserving a precious resource, more of it will be misused and wasted.

Developers have not been blind to regulatory dynamics. An article in the *San Diego BCA Builder* (Waite 1975, p. 10) advised its readers on how to cope with public proceedings: "It is generally advisable to be prepared to give something up. Don't go into the public hearing with a bare minimum proposal. Be ready to barter something away."

Many developers try to beat the game by asking for more than they really want, which frequently means the entire process accomplishes little more than wasting the public's time and money, increasing development costs, and further undermining the government's credibility.

Land Use Law

In two decisions issued in the past few years, the U.S. Supreme Court significantly limited zoning powers. *First English Lutheran Church v. County of Los Angeles*[15] held that the landowners are entitled to compensation when a regulatory ordinance denies them all use of the land, notwithstanding that the purpose of the law is to protect the public health and safety. Compensation is required even if the denial of use is only temporary. *Nollan v. California Coastal Commission*[16] held that for a land use regulation to be valid, it must substantially advance a legitimate state interest. The court has also ruled in recent years that a land use regulation can effect a taking if it denies owners the economically viable use of their land.[17] Together, these decisions place a burden on land use regulators to provide considerable justification for their rules, perhaps at times even on

[15]482 U.S. 304 (1987).

[16]483 U.S. 825 (1987).

[17]Agins v. Tuburon, 447 U.S. 255, 260 (1980); Penn Central Transportation Co. v. New York City, 438 U.S. 104, 127 (1978).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the order of a vital and pressing standard. It is clearly not enough to achieve regulatory validity for government merely to assert that it is implementing the police power.

In *Nollan*, the Court struck down the California Coastal Commission's requirement that the Nollans convey to the public an easement paralleling the coast. The Court reasoned that such an easement did not substantially advance the alleged state interest of making the coast accessible physically or visually to the public. "Substantially advance," according to the Court, means that an essential nexus exists between the means chosen to advance the end and the end itself; compliance with the standard will be carefully scrutinized—it is far "more than an exercise in cleverness and imagination."

The Personal Environment

Needlessly consuming land obviously limits the supply of that resource. Moreover, prohibiting or limiting development will injure the portion of the environment that is unquestionably the most important, that which houses the people and supplies their material needs.

For most people, the home is where the major part of life is spent. Its characteristics greatly influence the quality of one's life. Similarly, a better neighborhood and greater convenience to work and shopping are most significant environmental goals.

One direct consequence of these antidevelopment policies will be to deny to those people who would occupy the new and excluded housing an opportunity to better their environment. Another is that the living conditions of people who depend on the construction and related industries for their livelihoods will suffer temporarily if not permanently by the resultant curbs on production.

Frieden (1979) discusses three proposals to develop major housing projects in Northern California opposed by, among others, environmental groups. These projects were originally intended to provide more than 25,000 units, but the developers could obtain permission to erect only about 3,500. Even an environmentally enhanced development was sacrificed in favor of permitting only very low-density, upper-middle-class housing. Unit prices in the new developments were much higher than those originally contemplated, largely the result of the requirements imposed by governmental authorities.

Another consequence of antidevelopment policies was described in the *San Diego Union* (1988, B11). According to Herb Cawthorne, president and chief executive officer of the Urban League of San Diego, a proposed limited growth initiative then before the voters

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

would have brought "havoc to the working class people of San Diego." Bayard Rustin (1975, p. B3), long a leader in civil rights causes, observed that liberal-oriented environmentalists

> have failed to consider the implications of their creed that while a no-growth economy may protect the fields and streams (which in itself is a dubious claim), it will most certainly result in untold misery for thousands of ordinary people, many of whom are the black poor of America and the poverty-stricken masses of Asia, Africa and Latin America.

The indirect effects of antidevelopment policies will also be pervasive. Environmental conditions throughout a housing market will suffer. These practices will (a) curtail the filtering process upon which great numbers of people depend for better housing conditions, (b) generally raise the cost of shelter, (c) cause worse environmental conditions for many or most apartment dwellers within that housing market, and (d) cause greater spreading and scattering of housing accommodations, something popularly condemned as urban sprawl.

Filtering in housing occurs when new homes and apartments are constructed and families move into them, vacating their former residences for occupancy by others. The others, in turn, may vacate still other units and the process continues through a number of sequences. In the mid-1960s the Survey Research Center at the University of Michigan studied this process in 17 metropolitan areas. The results (Lansing et al. 1969) show that, on the average, construction of one new unit makes it possible for a succession of 3.5 moves to occur, each move being to different and probably better housing. New construction thus benefits more people indirectly than directly: 2.5 moves to existing housing and only 1.0 move to new housing.

The survey (Lansing et al. 1969, pp. 41, 68) showed that more than one-third of all people who move are likely to be in the lower- and moderate-income categories. Thus, for every 1,000 new housing units built, there are more than 3,500 relocations, of which 333, on average, are by poor families and 933 are by moderate-income families. The survey (pp. 19–20) also showed that while most construction occurs in the outer portions of a metropolitan area, those moves extend to older areas near the city's center. Because the nation's worst environmental conditions exist in these areas and for these lower-income levels, conditions will further deteriorate if filtering is reduced. In the United States, housing subsidy programs have done little to alleviate these problems; improvement necessitates a greater housing supply in the private market. For government

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

to elevate the living standards of 1,000 families would require enormous expenditures.[18]

Those of higher income who are ineligible for subsidized housing and who cannot afford new housing similarly depend for better housing on the filtering process induced by new construction. The University of Michigan survey makes it clear that prohibiting new construction seriously harms the groups that are in need of a better housing environment.

If government curtails construction in a housing market, prices will increase and competition will decrease.[19] In the case of apartments, rents will be raised either directly through the amount actually charged or indirectly through reduction in services, repairs, maintenance, and improvements (probably, in many instances, through a combination of all or most of these.[20] Higher rents will cause more doubling up and will increase density within areas containing these buildings. These conditions will be intensified as the filtering process is halted or reduced. The result will be much poorer housing conditions for many people and a further deterioration of marginal or older buildings and areas. Buildings will remain in service that otherwise would be removed.[21] More people will pay more for the same or lesser quality housing.

As I have previously observed, the antidevelopment pressures may curtail much construction, but they cannot stop it all. The demand

[18]I am told by builders that in 1989, the cost of construction of a 1,000-square-foot, moderately equipped apartment in San Diego was about $50,000. Accordingly, construction of 1,000 new units would cost $50 million plus land, which could be priced at $20,000 to $40,000 per unit in low- and moderate-income areas.

Some localities have sought to provide lower-cost housing by requiring developers to set aside portions of their developments (from 10 percent to 20 percent) for this purpose. Among other things, such requirements erode development feasibility, thereby limiting production of houses. For an analysis of such zoning restrictions—usually referred to as "inclusionary zoning," see Ellickson (1982, p. 135).

[19]See the discussion by Frech (1982, p. 259) about economic impact of the rules promulgated by the California Coastal Commission, which controls development along California's coast. For an interesting socioeconomic analysis of growth controls, see Construction Indus. Ass'n v. City of Petaluma, 375 F.Supp. 574 (N.D. Cal. 1974), *rev'd on other grounds*, 522 F.2d 897 (9th Cir. 1975), *cert. denied*, 424 U.S. 934 (1976). In this case, a federal judge found that the growth controls in Petaluma, California, would operate to limit housing supply and increase prices in the region.

[20]Increases in rental costs will also increase pressure on government to adopt rent controls, another regulation that limits production of housing and, therefore, raises its price. See Fraser Institute (1975).

[21]Opposition to zoning changes allowing construction of new apartment buildings may at times come from existing apartment owners who fear the influx of new competition. Quite often, business people seek government regulation to achieve personal advantage.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

for new housing will remain largely unabated; housing will be forced to locate in places of least resistance, where opposing political pressures are absent or limited. The result will be a greater proliferation or scattering of housing and of all problems that come with urban sprawl.

Similarly, density restrictions reducing yield per acre will cause housing to spread out. There restrictions will necessitate installing more streets and other facilities and creating greater burdens and inconveniences for more people. Obviously, a two-acre minimum lot size will generally require installing more lineal feet of road than a one-acre restriction. An apartment or townhouse project containing many more times the number of residents will require only a fraction of the lineal footage of pavement needed for a single-family development.

A study[22] of infrastructure costs for 1980 in Westchester County, New York, shows that as density decreases, these costs dramatically increase:

			Cluster
Lot Size	1 acre	½ acre	2 units/acre
Lot Width	125 feet	25 feet	35 feet
Infrastructure Costs	$30,125	$18,075	$8,435

Infrastructure costs include costs for streets, curbs, lighting, sidewalks, sanitary sewer, and water lines.

The Experience of Houston

Houston, Texas, is the only major city in the United States that has never adopted zoning.[23] Despite this, it is likely that land uses on the whole are about as separated in Houston as they would be under zoning. But even if this conclusion is not entirely accurate, it is more than offset by the economic and social rewards emanating from the

[22]See Urban Systems Research and Energy (1982, p. 71), *Report* (1982, p. 203, Table 15.1, n.20).

[23]Siegan (1972, p. 25) states that Houston is also the only major city that held straw votes on adopting zoning. In 1948, only property owners were allowed to vote, and the result was 14,142 to 6,555 against zoning. In 1962 when there was no such restriction on voting, the vote again went against zoning, 70,957 to 54,279; approximately 48.5 percent of those qualified voted. To date, no other vote has been held on this issue. Houstonians retain, of course, the option to impose zoning if they wish.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

absence of unnecessary government restraints over the development of property.[24]

How have these benefits occurred without zoning controls? The answer is that Houston does have land use controls, but they are primarily economic. Specifically, the use and development of land and property in Houston are controlled in three different ways: first, by the normal economic forces of the marketplace; second, through legal agreements, principally restrictive covenants; and third, through a relatively limited number of land use ordinances adopted by the city (Siegan 1972, pp. 26–27, 30–31).

Houston also controls development through subdivision, building, traffic, and housing regulations that do not appear to differ appreciably from those of other cities in its region. Most cities in the region use either the Southern Standard Building Code or the Uniform Building Code, with modifications. Both were drafted by private industry organizations. However, Bjornseth (1983, pp. 43–44) explains that Texas law prevents counties from adopting and enforcing building codes, and city governments' extraterritorial jurisdiction does not include any power to enforce building codes beyond city limits.

The contrast between Houston's controls and zoning is clear: Unless the property is subject to an enforceable restrictive covenant, the city exercises minimum control over the uses that will be made of and on the property.

The Houston experience demonstrates that, in the United States, zoning schemes regulating land use separation and yield are not essential to the existence and viability of cities. The real estate marketplace does not operate chaotically or haphazardly. Residential, commercial, and industrial uses tend to develop separately. Certain uses will develop only in certain places. Gas station and fast-food franchises, plus most other major commercial developments, provide an obvious illustration; regardless of where they are permitted, they will locate only on heavily trafficked streets.[25]

Thus, major business and commercial uses generally will be absent from residential or local streets that constitute 80 percent of total street mileage within Houston, and probably about the same elsewhere. In areas of Houston no longer subject to restrictive covenants

[24]See generally Siegan (1972) and Jones (1980). Jones was formerly director of city planning in Houston. According to Carreau (1981, p. 5A), when Jones appeared in 1981 before a public hearing of President Reagan's Commission on Housing, Jones said, "Zoning is archaic. It makes the marketplace respond to outmoded ideas."

[25]On the rare occasion when this does not occur, the life of the enterprise will be very short.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(to be discussed later) or in which covenants were never imposed, these local streets contain relatively few commercial uses, probably no more than 5 percent within a specified area. The bulk of these commercial uses are home occupations and businesses that serve area residents and, therefore, contrary to much conventional thinking on the subject, are probably compatible with the area.[26]

There is also a great tendency for industrial uses to group and concentrate separate from residential uses. A comparison of maps showing industrial locations in metropolitan Houston to similar maps of Los Angeles or Dallas (cities to which Houston is often compared) suggests that the proliferation of industry in Houston is probably no greater than in other metropolitan areas. There may even be less dispersal, although the large territories involved and the differing definitions of industry make measurement very difficult. This pattern is generally confirmed by land use maps of other nonzoned Texas cities: Pasadena, Wichita Falls, Laredo, and Baytown.[27]

It is generally too costly in terms of land prices and potential residential hostility for heavy industry to locate near new residential subdivisions (Siegan 1972, pp. 62–65). In Houston, plants and factories that are contiguous to and that were erected subsequent to homes are usually "light" rather than "heavy" in character. In most instances, their existence appears to pose no more, and possibly less, peril to residential values than would be the case if the same property had been developed for an alternative use, such as apartments.

Apartment and condominium development also reflects this pattern of separation. Thus, notwithstanding the absence of location restrictions, the vast bulk of multifamily development in Houston has occurred in the southwest section of the city.

[26]These conclusions are based on surveys (Siegan 1972, pp. 36–42) made in December 1969 of three substantial areas of Houston that had never been or were no longer restricted: Denver Harbor (950 structures), Montrose (450 structures), and Riverside (240 structures). Denver Harbor was never subject to restrictive covenants; most covenants in Montrose expired in 1936; and one-half of Riverside's covenants expired in 1950.

Houston contains a great many areas of single-family occupancy that look identical to zoned areas. There are also mixtures of uses available for those seeking to demonstrate the ill effects of nonzoning. Frequently, however, the same mixtures exist in cities that are zoned (Siegan 1970, pp. 71, 91 n.42.) Houston's appearance has long been a subject of considerable controversy. For a favorable view of this aspect of the city, see Huxtable (1976a, 1976b). At that time, Huxtable was the architecture critic of the *New York Times* and a Pulitzer Prize winner.

[27]See the following for maps: Bernard Johnson Engineers (1964, pt. 7, p. 32); Marmon, Mok & Green (1967, p. 77); Texas Highway Department (1964, fig. 9); Wichita Falls, Texas, (1964, fig. 9).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Substantial areas in and around Houston exhibit a minimum demand for multiple-family, industrial, and commercial development. These areas provide land for single-family occupancy. Most single-family developers in Houston (as well as in many other parts of the country before the advent of zoning) have traditionally imposed restrictive covenants to permit building only houses with specified characteristics within their subdivisions.[28] Because many of the earlier restrictive covenants in Houston were limited in duration or were legally insufficient or not enforced by owners, zoning would probably have kept more areas for single-family occupancy (Siegan 1972, p. 33). As it happens, zoning would thereby have impeded development both of much housing and of nonhousing facilities, which did occur subsequently within the city.

Studies of zoning have shown that market mechanisms reduce the impact of uses that are regarded in zoning theory as adverse to property values (Crecine, Davis, and Jackson 1967, p. 79; Maser, Riker, and Rosett 1977, p. 111; Reuter 1973, p. 313; Stull 1975, p. 535). The same phenomenon is evident in Houston where, of course, no zoning ordinance is available to advise purchasers as to potential use of adjoining land. Typically, land on the perimeter of a residential subdivision will sell for far less than that located in the interior when the land adjoining the subdivision is vacant or used for purposes other than residential development. In a nonzoned market, economic forces operate to internalize many externalities of land development and use.

Most covenants created subsequent to World War II are much more durable than those in the past and seem to offer a reasonably practical solution to the conflicting desires of allowing for change and yet maintaining stability. Most post–World War II covenants contain an automatic extension provision (Siegan 1972, pp. 34–35). They provide for an initial duration period of 25–30 years, and an indefinite number of 10-year automatic extension periods. Agreement on the part of 51 percent of the owners (usually one vote per lot or on the basis of frontage) may cancel or amend the covenants before the end of the initial period or before the end of any subsequent 10-year period. Under this provision, a majority of homeowners can control the destiny of their subdivision.

[28]In contemporary years, regardless of whether zoning exists in the community, developers of planned unit developments—both detached and attached housing—usually impose restrictive covenants as a sales device to accommodate the needs and desires of the residents.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Inasmuch as enforcement of restrictive covenants can be costly for homeowners in lesser-income subdivisions and small subdivisions, Houston adopted an ordinance in 1965 enabling the city to enforce these covenants.[29] Houston has also adopted an offstreet parking ordinance for residential development, a limited number of location restrictions, and a relatively small number of other government regulations designed to cure problems of land use not satisfactorily controlled by the private market.[30]

For homeowners, restrictive covenants serve the same purpose of maintaining exclusivity as does zoning. While similar in this respect, the covenants otherwise vary greatly from zoning both in application and operation; they illustrate the difference between the economic and political marketplaces in determining land use and development.

As heretofore explained, zoning is controlled by the political system and principally achieves that which is most important politically. It allows homeowners to influence zoning of land that is far removed from their subdivision. On the other hand, developers or owners and their lenders impose covenants only on their subdivisions as a means to secure and maximize their investments. They will apply covenants in accordance with what they believe are the desires of their prospective purchasers. Since there is usually no incentive for owners to restrict the use of their land while it is in a raw state, covenants normally affect little more than land already developed or programmed for development, and then largely for homes or row houses. As a result, probably no more than 15–20 percent of the land in Houston is subject to restrictive covenants. Under zoning, every square inch of the city's land would be regulated.

Restrictive covenants are a market device to maximize property value. Most American homeowners prefer to live in a homogeneous environment, and they should have the freedom to pursue this goal, provided others are not harmed. Restrictive covenants come close to achieving this balance. By contrast, zoning allows almost unlimited pursuit of exclusivity, often with adverse effects on many.

[29]In early 1990, some community and political leaders in Houston called for the adoption of zoning, charging in part that the city had been lax in enforcing restrictive covenants. I believe there is substance in these charges. In addition, these accusations spread fear among homeowners and may lead to the impositions of new land use regulations. For information on the city's power to enforce the covenants, see Susman (1966, p. 741); Baylor Law Review (1969, p. 307); Houston Law Review (1972, p. 816).

[30]Ordinances as of 1972 are set forth in Siegan, (1972, pp. 26–31). Additional regulations are reported in *New York Times* (1982, p. A12). The city has also in recent years passed ordinances prohibiting "undesirable" uses such as sexually oriented businesses, junk yards, and helicopter pads in or near residential areas.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Houston Compared to Dallas

When one wants to quantify the results of zoning, it is useful to consider the experience of Houston and its environs. By comparing Houston with a similar area that has imposed zoning, one can estimate the consequences of regulation. The most obvious area for comparison is Dallas. Houston and Dallas are approximately 230 miles apart and share similar economic profiles.[31]

Dallas has been zoned since the early 1930s. The bulk of the Houston Standard Metropolitan Statistical Area (SMSA) is not zoned, while approximately the reverse situation occurs in the Dallas SMSA. Zoning in Dallas is of a relatively moderate character. It is far less severe than zoning regulation in many of the nation's developing suburbs, including some Houston and Dallas suburbs that have enacted strict zoning control. The other regulations affecting housing in the two cities (e.g., subdivision, building, minimum housing, and traffic) do not differ appreciably.[32]

Until the fall of world oil prices in the early 1980s, both areas had for many years experienced a very high housing demand. There is no evidence that the spectrum of this demand between the two areas varied significantly. The similarity of economic statistics over the years supports this conclusion. However, the supply response in each differed considerably; the Houston area catered much more to average- and lesser-income people by producing a far greater percentage of multifamily units. Land use regulation appears to be the major factor accounting for this difference since its presence in Dallas likely restricted apartment production there. As a general

[31]For comparisons of Houston and Dallas over the 1960–70 period, see Siegan (1972, pp. 100–3). Similar relationships appear in the 1980 census figures (U.S. Department of Commerce 1980b, pp. 45–21.) One difference is that the median income in 1979 was about 10 percent higher in the Houston SMSA than in the Dallas–Fort Worth SMSA. This difference does not affect any Houston-Dallas comparisons contained in this article.

[32]In recent years, most building in the Houston area has occurred in the unincorporated sections of Harris County, which has virtually no building regulations. I am uncertain as to this situation in the Dallas area. The absence of building codes would operate to reduce cost of construction. Accordingly, the comparisons between Houston and Dallas are subject to possible modification, depending on how much new construction has occurred in areas without building codes. Houston's building costs are apparently higher than those in Dallas (U.S. Department of Labor 1972–79). Boeckh Building costs show that for the sample years of 1970, 1976, and 1980, Houston construction costs were a little higher than those in Dallas (Boeckh Division 1982). Because of Houston's favorable attitude toward development, the difference between building costs in Houston and unincorporated areas of Harris County are probably in the low single-digit area. As of 1980 Houston used the 1970 Uniform Building Code, and Dallas used the 1979 Uniform Building Code, with each imposing its own modifications.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

matter, zoning is much more restrictive toward multiple-family than single-family dwellings.

The comparisons below are between Harris County, which contains Houston, and the Dallas SMSA. Harris County's population was 12 percent less in 1970 and 17 percent less in 1980 than that of the Houston SMSA. About three-quarters of the SMSA's housing starts occurred in that county. This comparison uses Harris County and not the Houston SMSA because the figures come from the same source as those for Dallas SMSA; this source, unlike others, separates townhouses and condominiums, which, for purposes of this analysis, should be regarded as multifamily housing (Siegan 1983, pp. 720, 728 n.1). This source does not provide statistics for the Houston SMSA.

The portion of the Houston SMSA that is excluded by using Harris County figures produces about 50 percent multifamily housing; therefore, the SMSA would show somewhat less difference from the Dallas SMSA in the category of construction. Multifamily housing in Tables 1 and 2, accordingly, includes townhouses (not identified as

TABLE 1

MULTIFAMILY AND SINGLE-FAMILY HOUSING STARTS IN HOUSTON AND DALLAS AREAS, 1973–80

	Harris County	Dallas SMSA	Harris County % Higher
Total Housing Starts	300,732	173,854	73
Total Multifamily Starts	173,158	80,266	116
Total Single-Family Starts	127,574	93,588	36
Population Growth, 1970–80	653,488	393,050	66
Individual Housing Units, 1970	587,671	529,785	11
Population, 1970	1,741,912	1,555,950	12
Population, 1980	2,395,400	1,949,000	23

TABLE 2

PERCENTAGE OF MULTIFAMILY AND SINGLE-FAMILY HOUSING STARTS FOR HOUSTON AND DALLAS AREAS, 1973–80

	Harris County (%)	Dallas SMSA (%)
Multifamily Starts	58	46
Single-Family Starts	42	54

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

to rental or sale) and condominiums. (Tables 1 and 2 are reproduced from Siegan 1983, p. 728.) Inasmuch as multifamily housing occupies much less land than single-family dwellings, Houston has used considerably less land than Dallas for housing purposes.[33]

This difference in multifamily output probably accounts for rents in Houston being lower during the 1973–80 period than in Dallas. The Bureau of Labor Statistics (U.S. Dept. of Labor 1972–79) provides the most comprehensive rent comparisons for the two areas in its annual estimates of living costs at three income levels for a four-person family and a retired couple. For 1972–79, inclusive, these estimates show that Houston SMSA rents were the same or lower in all categories. Totaling all these figures on an unweighted basis, the Dallas SMSA rents were about 13 percent higher than Houston SMSA rents. For the other years for which comparable figures are available for both SMSAs, 1966–69, inclusive, the difference in rental prices was about 15 percent (Siegan 1972, p. 112).

Support for these findings comes from the *Eton Journal* surveys of 1979, 1980, and 1981 apartment rentals in the nation's 16 most significant real estate markets. For units located in the city, Dallas rents were 14 percent higher than Houston's in December 1980 and 16 percent higher in March 1981. For units identified as "suburban," Dallas area rents were 17 percent higher in December 1980, 25 percent higher in March 1981, and 11 percent higher in May 1981. In January 1979, Dallas rents were 5 percent higher in both city and suburban sections and 12 percent higher in both sections in January 1980 (Siegan 1983, p. 729).[34]

[33]The actual housing productivity of the Houston area may be far greater than the foregoing figures suggest. According to Chicago Title and Insurance Company Market Research Services, Chicago, Illinois, in 1978–81 the Houston SMSA produced many more new residential units per 100 existing households than the Dallas–Fort Worth SMSA or the U.S. average. Chicago Title (1980) and Pfister (1982, p. 12) show the following units per households:

	Houston SMSA	Dallas–Ft. Worth SMSA	U.S. Average
1978	7.35	5.02	2.61
1979	5.68	4.52	2.19
1980	3.58	3.05	1.64
1981	4.39	2.62	1.35

Figures for 1980 and 1981 are reported in Pfister (1982, p. 12) and figures for 1978 and 1979 are contained in Chicago Title (1980). The 1966–69 figures are in Siegan (1972, p. 112).

[34]Notwithstanding these statistics, the 1980 census shows that the median contract rent for Houston SMSA was $256 and for Dallas–Fort Worth SMSA, $239; rent was $258 and $244, respectively, for the urbanized areas (U.S. Department of Commerce 1980a, pp. 45–9 to 45–10). The evidence that has come to my attention and herein reported

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Moreover, throughout this period, vacancy rates in Houston apparently were much higher than those in Dallas. According to the *Eton Journal* figures, Houston averaged the highest vacancy rate (6–7 percent) of any city surveyed in both January 1979 and 1980 (Siegan 1983, p. 729).[35] This finding is not surprising, because Houston frequently has had large vacancy rates resulting from its high productivity.

Comparable price statistics for single-family dwellings do not appear to be available. Available home prices show a different pattern from apartment rentals. The Federal Housing Administration (FHA 1960–75) has annually published prices and other data relative to sales of existing and new homes that it has insured under its Section 203 program. This program principally involves the most inexpensive nonsubsidized housing available in the market. According to these surveys, from 1960 through 1974 there was often little difference in actual prices (taking into account construction, square footage, and amenities) of existing or new FHA-insured homes between the Houston and Dallas SMSAs. For most of this period, Houston's prices were lower, although in some years they were noticeably higher.

Beginning in 1975, FHA changed the measurement area from Dallas SMSA to Dallas–Fort Worth SMSA, of which Dallas is a much smaller component, limiting the relevance of this information. Thus, the population of Dallas in 1980 was 904,078, while that of the Dallas SMSA was 1,949,000 and the Dallas–Forth Worth SMSA was 2,960,000. For 1980, Houston's population was 1,594,086 and that of the Houston SMSA was 2,887,000 (Siegan 1983, p. 730 n.11).[36]

I am aware of only one study concerning residential price differences in Dallas and Houston. Richard Peiser (1982, p. 30) observed that when similar houses in equivalent locations were compared, prices were about $4,000 to $10,000 less expensive in Houston. He attributes this difference largely to the higher price of land in Dallas, which was brought about by zoning and other regulations in that city.

A study at the Urban Land Institute (Dunau 1981) determined land prices and rates of their increase between 1975 and 1980 for 30

---

leads me to discount the 1980 census figures as revealing the actual differences in rent between the two cities.

[35]The 1980 census (U.S. Department of Commerce 1980a, pp. 45–9 to 45–10) shows the rental vacancy rate for the Houston and Dallas–Fort Worth SMSAs as 14.7 percent and 10.4 percent, respectively. For the urbanized areas, the figures were 14.5 percent and 10.5 percent.

[36]The original figures are taken from FHA (1960–75) annual editions and Bureau of Census (1981).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 3

LAND PRICES AND RATES OF INCREASE, 1975–80

SMSA	Improved Lot 1980	Improved Lot 1975	5 yr. % Increase	Raw Acreage 1980	Raw Acreage 1975	5 yr. % Increase
Dallas	$16,000	$9,500	68.4	$20,500	$8,500	141.2
Houston	12,000	7,850	53.8	10,000	7,004	42.8

SMSAs. Table 3 shows the prices and percentage of increases for Dallas and Houston SMSAs.

Housing cost generally should be cheaper in the absence of regulation. President Reagan's Housing Commission (*Report* 1982, p. 180) concludes that unnecessary zoning and related requirements may often elevate the price of housing by 25 percent or more.

Economics professors Lloyd Mercer and W. Douglas Morgan (1982, p. 189) conducted a study of Santa Barbara County Housing and found that development restrictions accounted for more than 27 percent of the increase in real housing prices during 1972–1979.

Sogalyn and Sternlieb (1972) studied certain zoning and building requirements in New Jersey. These authors concluded that reducing three major zoning requirements (lot size, lot frontage, and living area) would reduce prices considerably and enlarge the effective housing market. Changing two building code specifications (thickness of exterior wall sheeting and size of foundation cinderblock) would also lower selling prices, but not to the same degree as would altering zoning policies.

To determine the impact of local government regulation on the cost of housing, the Department of Housing and Urban Development initiated a Housing Cost Reduction Demonstration project in 1980. Four communities across the country were selected to participate in the project, which used reduced local government regulations as the only variable. In these communities, zoning, building, and subdivision regulations were limited. In the initial projects for which early figures were available, the prices of homes were reduced by 21 percent to 33 percent. In Shreveport, Louisiana, demonstration housing units had a sales price of $52,850 while homes in a comparable suburban project with conventional regulations and processing sold for $70,000. In Hayward, California, the demonstration units ranged in price from $53,000 to $65,000. Comparable units in the area sold for $79,500 to $97,500. In all instances, the builders sought to obtain a normal profit margin (Bjornseth 1983, p. 44).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Low-Cost Housing

Zoning rules have prohibited the construction of many low-cost accommodations. Again, Houston provides an interesting example. The city has no density controls, and in the late 1960s projects were built in minority neighborhoods containing 16 or more detached units.[37] This density was even too much for Houston's planning department, which sought unsuccessfully to ban them.

The developments in question were being erected in very low-income black areas by private, unsubsidized investors (a rare occurrence in the United States). The principal inducement for the developers was the opportunity afforded by the large density to reduce land and construction costs. These projects offered unusually inexpensive rental housing, which was, in all likelihood, superior to that which the tenants last occupied.

Because Houston imposes no restriction as to size and type, the market for new houses is very flexible. As mortgage rates rise, for example, the developer can reduce the size of units to limit prices. In many suburbs of Dallas, on the other hand, where zoning regulates the size and type of dwelling, it is not possible to construct units below a specified square footage. A Texas developer testified before the Reagan Housing Commission that, as a result, he had to discontinue building in these suburbs, thereby denying many potential consumers an opportunity to purchase housing (*Report* 1982, p. 204).

Other Considerations

Critics of Houston's land-use system contend that, as a result of it, the city is not very pleasing aesthetically. Beauty being in the eye of the beholder, it is not difficult to find others who disagree and consider the city quite attractive. However, the use of land should not be judged by this or any other single factor. If aesthetics were the major concern, the tightly zoned, strongly exclusionary suburbs would be ideal. The issue is, instead, to determine the land-use system that bests serves society, an inquiry that involves the extensive analysis seen in this article.

One virtue of the Houston system that has been previously referred to is worthy of additional observations. The absence of regulation in that city affords great opportunity for builders and developers to satisfy consumer demand. Human resourcefulness and inventiveness thrive in Houston because of the absence of their enemy, gov-

[37]At that time, single-family detached units typically ranged in density in the United States between one and five units per acre (Council on Environmental Quality 1974).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ernment regulation. Unfortunately, in zoned cities, these talents are often spent in persuading or outmaneuvering the zoning authorities.

The communist collapse reveals how debilitating government regulation is to human creativity, ingenuity, and productivity. Communist rulers sought to regulate almost every human activity to achieve the common good; instead of producing a paradise, however, they created virtual prisons. These prisons did not necessarily confine physical bodies, but they did incarcerate the human mind.

Conclusion

Zoning practices seriously burden the use and development of land. Consider the fate of the last 40 acres of vacant land that remain in a zoned community where all other land has been fully developed. Assume the zoning on the acreage has not been finalized (that is, it is in a holding classification). The decision as to how this tract will be used in a zoned city will be made through the planning and political processes of local government. The city council or other governing board will classify the property after hearing from all interested parties and their planners, lawyers, and other representatives. These parties include owners, neighbors, local civic or homeowners' groups, political organizations, school and park boards, labor groups, perhaps the Chamber of Commerce, and, of course, assorted do-gooders and do-badders. Almost everyone except those who will directly benefit from the development, such as potential homeowners, tenants, and shoppers, can be expected to enter the fray. Each side will have no difficulty in producing a planner to prove "conclusively" that its position is the only correct one.

How the debate will come out will depend on who or what is best able to influence or pressure or even pay for the vote of city council members. That there may be an enormous demand for apartments in the community is likely to be a less-important factor than the opposition to such use by certain politicians, homeowners, civic groups, or the media.

By contrast, the answer where there is no zoning is relatively clear. The property will, in all likelihood, be developed for the most valuable use—the one that society values the greatest. The land will, thereby, satisfy the predominant consumer demand. This example assumes, of course, that the city will still exercise controls under the vital and pressing standard previously discussed.[38] Thus, as the

[38]It also assumes that in a zoned area the prudent investment expectations of owners of adjoining and nearby property will be protected. Consider in this respect, the following recommendation of Reagan's Housing Commission (*Report* 1982, p. 201):

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Houston and Dallas comparisons reveal, a high demand for apartments is much more likely to be satisfied in the absence than in the presence of zoning.

The latter alternative will provide, I submit, the most socially rewarding use of the property as determined by the least fallible of city planners, the marketplace. It will provide more material and environmental benefits for more people, both those who benefit directly and those who benefit indirectly from the project. It will make for a more-efficient allocation of resources within the community. And precious land will be used to provide for the needs and desires of human beings and will not be wasted for the sake of political expediency and planners' speculations.

References

"An Interview with Milton Friedman." *Fraser Forum* (July 1989): 7.

Aron, Leon. "What Glasnost Has Destroyed." *Commentary* (November 1989): 30–31.

Baylor Law Review. "Comment. Houston Invention of Necessity: An Unconstitutional Substitute in Zoning." *Baylor Law Review* 21 (1969): 307.

Bernard Johnson Engineers, Inc. *Summary Report of the Comprehensive City Plan, Baytown, Texas,* 1964.

Bjornseth, Dick. "No-Code Comfort." *Reason* (July 1983): 43–44.

Boeckh Division, American Appraisal Company. *Boeckh Building Cost-Index Numbers,* January-February 1982.

Brzezinski, Zbigniew. *The Grand Failure: The Birth and Death of Communism in the Twentieth Century.* New York: Charles Scribner's Sons, 1989.

Buchanan, James M. "The Constitution of Economic Policy." *American Economic Review* 77 (June 1987): 243–50.

Buchanan, James M., and Tullock, Gordon. *The Calculus of Consent: Logical Foundations of Constitutional Democracy.* Ann Arbor: University of Michigan Press, 1962.

Bureau of the Census. *Advance Reports, 1980 Census of Population in Houston, Texas.* Washington, D.C.: Government Printing Office, 1981.

Burka, Paul. "Liberal Education." *Texas Monthly* (March 1977): 109.

Carreau, Mark. "Hearings on Housing, City Planner Tells Commission Zoning Archaic." *Houston Post,* 15 December 1981.

A possible problem of deregulation [of zoning] is that it may adversely affect those who in good faith made their purchase or investments in reliance on the old rules [the prior zoning ordinance]. A change to the proposed "vital and pressing" standard would pose such a problem. Persons who purchase a home or a lot for construction of a home near vacant land assume that it will not be arbitrarily reclassified to allow other uses. The reasonable investment expectations of these homeowners should be protected. When vacant land is proposed for a use that would have required rezoning, homeowners entitled to notice under the old rules should be protected under the requirements and procedures of the old rules.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Chicago Title Insurance Company Market Research Services. *1980 New Residential Units per 100 Households Forecast,* 1980.

Coase, Ronald. "Economists and Public Policy." *In Large Corporations in a Changing Society.* Edited by F. Weston. New York: New York University Press, 1975.

Council of Economic Advisers. *Economic Indicators* 24 (December 1989).

Council on Environmental Quality. *The Costs of Sprawl.* Washington, D.C.: Government Printing Office, 1974.

Crecine, John P.; Davis, Otto A; and Jackson, John E. "Urban Property Markets: Some Empirical Results and Their Implications for Municipal Zoning." *Journal of Law and Economics* (1967): 79–99.

Dunau, Frank R. "Land Price Levels and Rates of Increase." Working Paper, Urban Land Institute, Washington, D.C., 1981.

Economic Report of the President. Washington, D.C.: Government Printing Office, February 1990.

Ellickson, Robert. "The Story of 'Inclusionary Zoning.'" In *Resolving the Housing Crisis,* pp. 135–87. Edited by M. Bruce Johnson. Cambridge, Mass.: Ballinger Publishing Co. for the Pacific Institute, 1982.

Federal Housing Administration. *Data for States and Selected Areas of Characteristics of FHA Operations under Section 203.* Washington, D.C.: Government Printing Office, 1960–75.

Fraser Institute. *Rent Control: A Popular Paradox,* Vancouver, B.C., 1975.

Frech, H. E. III. "The California Coastal Commissions: Economic Impacts." In *Resolving the Housing Crisis,* pp. 259–74. Edited by M. Bruce Johnson. Cambridge, Mass.: Ballinger Publishing Co. for the Pacific Institute, 1982.

Frieden, Bernard. *The Environmental Protection Hustle,* Cambridge: MIT Press, 1979.

Gorbachev, Mikhail. *Toward a Better World.* New York: Richardson and Steirman, 1987.

Houston Law Review. "Comment. The Municipal Enforcement of Deed Restrictions: An Alternative to Zoning." *Houston Law Review* 9 (1972): 816–84.

Huxtable, Ada Louis. "Houston Is the Future." *Houston Chronicle,* 22 February 1976a.

Huxtable. Ada Louis. "Houston Will Never Be New York." *Houston Chronicle,* 22 February 1976b.

Johnson, M. Bruce. "Piracy on the California Coast." *Reason* (July 1974): 18.

Jones, Roscoe. "Houston City Planning Without Zoning." In *Zoning, Its Costs and Relevance for the 1980s.* Edited by M. Goldberg and P. Horwood. Vancouver, B.C.: The Frazer Institute, 1980.

Lansing, John B.; Clifton, Charles Wade; and Morgan, James N. *New Homes and Poor People: A Study of the Chain of Moves.* Ann Arbor: University of Michigan Institute for Social Reason, 1969.

Marmon, Mok & Green, Inc. *Development Plan for Pasadena, Texas,* 1967.

Maser, Steven A.; Riker, William H.; and Rosett, Richard N. "The Effects of Zoning and Externalities on the Price of Land: An Empirical Analysis of Monroe County, New York." *Journal of Law and Economics* 20 (1977): 111–32.

Mercer, Lloyd, and Morgan, W. Douglas. "An Estimate of Residential Growth Controls Impact on Housing Prices." In *Resolving the Housing*

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Crisis, pp. 189–215. Edited by M. Bruce Johnson. Cambridge, Mass.: Ballinger Publishing Co. for the Pacific Institute, 1982.

National Commission on Urban Problems to the Congress and to the President. *Building the American City*. Washington, D.C.: Government Printing Office, 1968.

New York Times. "Houston's Council Tries to Control City Growth." *New York Times*, 23 June 1982.

Peiser, Richard. "Lot Costs Compared in Dallas and Houston." *Tierra Grande* 18 (1982): 30.

Pfister, Richard. "The Housing Recession." *The Guarantor*, May-June 1982.

President's Commission on Urban Housing. *A Decent Home*. Washington, D.C.: Government Printing Office, 1969.

Report of the President's Commission on Housing. Washington, D.C.: Government Printing Office, 1982.

Reuter, Frederick H. "Externalities of Urban Property Markets: An Imperical Test of the Zoning Ordinance of Pittsburgh." *Journal of Law and Economics* 16 (1973): 313–49.

Reynolds, A. "International Comparisons of Taxes and Government Spending," Presented at a conference on "Rating Economic Freedom," Banff, Alberta, Canada, 5–8 October 1989.

Rustin, Bayard. "No-Growth and the Poor." *San Diego Evening Tribune*, 31 July 1975.

San Diego Union. "No-Growth Racism Is an Abuse," 26 August 1988.

San Diego Union. "Soviet Is Upset by Food," 16 October 1989.

Siegan, Bernard H. "Non-Zoning in Houston." *Journal of Law and Economics* 13 (1970): 71–147.

Siegan, Bernard H. *Land Use Without Zoning*. Lexington, Mass.: D.C. Heath, 1972.

Siegan, Bernard H. "Commentary on Redistribution of Income Through Regulation in Housing." *Emory Law Journal* 32 (1983): 720–27.

Shelton, Judy. *The Coming Soviet Crash*. New York: The Free Press, 1989.

Sogalyn, Lynne B., and Sternlieb, George. *Zoning and Housing Costs: The Impact of Land-Use Controls on Housing Price*. New Brunswick, N.J.: Rutgers University Press, 1972.

Stull, William J. "Community Environment, Zoning, and the Market Value of Single-Family Homes." *Journal of Law and Economics* 18 (1975): 535–57.

Susman, Thomas M. "Municipal Enforcement of Private Restrictive Covenants: An Innovation in Land-Use Control." *Texas Law Review* 44 (1966): 741–68.

Texas Highway Department. *Laredo Urban Transportation Study*, 1964.

Toll, S. I. *Zoned American*. New York: Grossman Publishers, 1969.

Urban Systems Research and Energy, Inc. *The Effects of Environmental Regulations on Housing Costs*, 1982.

U.S. Department of Commerce. *1980 Census of Housing (General Housing Characteristics—Texas)*. Washington, D.C.: Government Printing Office, 1980a.

U.S. Department of Commerce. *1980 Census of Housing (General Social & Economic Characteristics—Texas)*. Washington, D.C.: Government Printing Office, 1980b.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

U.S. Department of Labor. *Bureau of Labor Statistics Bulletin. Three Budgets for a Four-Person Family and for an Urban Retired Couple.* Washington, D.C.: Government Printing Office, annual editions, 1972–79.

Waite, Harrison. "How to Preserve an Endangered Species." *San Diego BCA Builder,* October 1975.

Wichita Falls Urban Transportation Plan, 1964–85, Wichita Falls: Texas Highway Department, 1964.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

A Positive Proposal to Improve Our Schools

James D. Gwartney

Education accounts for approximately one-third of state and local government spending. Perhaps no other item exerts a greater impact on our quality of life and economic future. If students are developing skills, acquiring knowledge, and preparing themselves for life in the modern world, our economic future is bright. On the other hand, declining basic skill levels are a prelude to economic stagnation (Bishop 1989). Thus, success in education is vitally important to every American.

The Education Record—We Are Spending More; Students Are Learning Less

Americans know that education is important. As a result, they have consistently supported increases in educational spending. Instructional expenditures per pupil in the United States exceed the levels of other industrial nations. In 1985, U.S. expenditures per pupil were 47 percent greater than in West Germany, 66 percent greater than in France and Australia, 74 percent greater than in the United Kingdom, and 83 greater than in Japan.[1] Unfortunately, the basic math and science skills of high school students in each of these countries exceed the level of U.S. high school students. In fact, U.S. high school students not only rank below students in Japan, West Germany, and the United Kingdom, they also rank below their counterparts in Spain, Ireland, and South Korea.[2]

During the past two decades the real expenditures per elementary and secondary pupil in the United States have increased substan-

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at Florida State University.

[1]U.S. Department of Education. Cited in Hood (1990).

[2]See United Nations Education, Scientific, and Cultural Organization (1988) and Lapointe, Mead, and Phillips (1989) for data on this topic.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tially. Measured in 1982–84 dollars, per pupil expenditures rose from $1,704 in 1967 to $3,501 in 1987, an increase of 105 percent (see Table 1). Because these figures are adjusted for inflation, they indicate that the actual purchasing power per pupil available to schools more than doubled during the 20-year period.

What impact has this vast increase in purchasing power for schooling resources had on student performance? As Table 1 illustrates, the achievement scores of college-bound high school seniors on both the Scholastic Achievement Test (SAT) and the American College Test (ACT) declined during the period. During 1967–80, the average SAT score of college-bound seniors in the United States went from 958 to 890, a 68-point decline. Since 1980, about 20 percent (14 points) of that decline has been regained. However, the SAT performance of high school seniors in the late 1980s remained well below

TABLE 1

REAL EXPENDITURES PER PUPIL ON ELEMENTARY AND SECONDARY SCHOOLING AND THE AVERAGE ACHIEVEMENT SCORES OF COLLEGE-BOUND HIGH SCHOOL SENIORS, 1967–88

	Real Current Expenditures per Pupil (1982–84 dollars)	Average Achievement Score of College-Bound Seniors	
		SAT	ACT
1967	$1,704	958	—
1969	1,896	956	—
1970	2,103	948	19.9
1972	2,368	937	—
1974	2,448	924	—
1976	2,643	903	18.3
1978	2,796	897	18.5
1980	2,757	890	18.5
1982	2,825	893	18.4
1984	3,054	897	18.5
1986	3,427	906	18.8
1987	3,501	906	18.7
1988	3,573	904	18.8
Percent Change:			
1970–88	+69.9	−4.6	−5.5

SOURCE: Center for Educational Statistics, *Digest of Educational Statistics, 1989*, pp. 120, 156. The expenditures are current expenditures per pupil in average daily attendance. The Consumer Price Index (1982–84 = 100) was used to convert the expenditure figures into constant dollars.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the levels of the late 1960s. The ACT scores of college-bound students have followed a similar path. As in the case of the SAT, the performance level of high school seniors on the ACT remains well below the average score achieved in 1970.

Puzzled by the decline in student performance accompanying the educational spending increases of the past two decades, researchers have tried to figure out what we are doing wrong. Many speculated that the schools were simply expanding resources in the wrong areas. According to this view, once the experts figure out the "educational production function" that works best, better results will be achieved.

However, recent research provides little comfort for those who believe that the situation will improve once we get the educational production function right. A monumental survey article by Eric Hanushek (1989) sheds light on this topic. Of 152 studies that examined whether a lower pupil to teacher ratio affected student performance, only 9 percent found a positive impact, while 8 percent found a negative impact. Of 113 studies that examined whether the education level of teachers makes a difference, only 8 (7 percent) found that it did. Less than 30 percent of the studies surveyed by Hanushek found a positive link between teacher experience and student performance. With regard to teacher salaries, only 16 percent of the studies found a positive relationship between higher teacher salaries and student performance, while 84 percent of the studies found that the higher salaries exerted no impact (or a negative impact) on student learning.

These findings imply a weak relationship (at best) between educational expenditures and student performance. Given the current structure of our educational system, increases in educational spending generally fail to improve student performance. Tables 2, 3, and 4 illustrate this point.

Under our federal system, the primary responsibility for education resides with the states. In recent years, some states have increased educational expenditures much more rapidly than others. Table 2 represents data for both the 10 states that increased *real* per pupil expenditures by the largest amount during 1970–88 and the 10 states that increased *real* expenditures by the smallest amounts. In 1970 the average per pupil expenditure levels of these two groups were approximately equal ($2,354 compared to $2,303). However, during the subsequent 18 years the two groups followed quite a different course. The first group of 10 states increased expenditures on elementary and secondary education quite rapidly. By 1988 the *real* per pupil expenditures of this group of 10 states had risen, on average, from $2,354 to $5,239, an increase of 122.6 percent. On the other

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 2

THE 10 STATES WITH THE LARGEST (AND SMALLEST) INCREASE IN REAL EXPENDITURES PER PUPIL, 1970–88

Category	Per Pupil Expenditures in 1988 Dollars		Percent Increase
	1970[a]	1988[a]	
10 States with the Largest Increase in Real Expenditures			
Massachusetts	$2,296 (21)	$5,466 (6)	138.1
Connecticut	2,689 (10)	6,396 (4)	137.9
New Jersey	2,936 (3)	6,933 (2)	136.1
Nebraska	1,607 (48)	3,600 (37)	124.0
Wyoming	2,470 (17)	5,503 (5)	122.8
Alaska	3,302 (2)	7,038 (1)	113.1
Delaware	2,418 (19)	5,138 (8)	112.5
Arizona	1,628 (46)	3,451 (38)	112.0
Maine	2,089 (30)	4,427 (15)	111.9
Virginia	2,107 (29)	4,436 (14)	110.5
Average	$2,354	$5,239	122.6
10 States with the Smallest Increase in Real Expenditures			
California	$2,811 (5)	$3,916 (26)	39.3
Utah	1,829 (40)	2,632 (49)	43.9
Idaho	1,918 (33)	2,778 (46)	44.8
Hawaii	2,595 (14)	3,859 (28)	48.7
Iowa	2,714 (8)	4,159 (22)	53.2
South Dakota	2,003 (32)	3,173 (41)	58.4
Arkansas	1,628 (46)	2,618 (50)	60.8
Montana	2,506 (16)	4,061 (24)	62.1
Minnesota	2,692 (9)	4,390 (16)	63.1
Nevada	2,329 (20)	3,800 (30)	63.2
Average	$2,303	$3,539	53.7

[a]Number in parentheses indicates the ranking of the state among the 50 states.
SOURCE: National Education Association, *Estimates of School Statistics.*

hand, real per pupil expenditures for the second group rose much more slowly—from $2,303 to $3,539, an increase of 53.7 percent. By 1988, the per pupil spending of the 10 states that increased spending most rapidly was 48 percent higher than the parallel figure for the 10 states with slower increases in educational expenditures.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

What did the states that increased educational expenditures rapidly get for their investment? Tables 3 and 4 shed light on this question. Table 3 presents data on the adjusted graduation rate, the proportion of ninth graders who graduated from high school three years later adjusted for interstate migration. By the 1980s, the per pupil expenditures were substantially higher in the states that increased educational expenditures rapidly. Nonetheless, the graduation rate of the states with the slower growth rate of per pupil educational expenditures was higher in both 1982 and 1987. Between 1982 and 1987, the graduation rate rose slightly more (4.7 percent compared to 3.3 percent) in the high-expenditures states. However, the increase was entirely the result of an exceptionally large increase in the graduation rate in one state (Wyoming). If Wyoming is deleted from the calculation, the increase in the graduation rate for the two state groups is virtually identical.

The graduation rate does not tell us anything about student learning. Achievement scores provide a better measure in this area. Table 4 provides data on SAT and ACT scores (relative to the national average) for the two state groupings. There is no evidence that the achievement scores in the "big educational spending" states are either higher or increased more rapidly in the 1980s. In fact, the average achievement scores in both 1982 and 1988 are slightly higher in the states with the slower growth rate of educational expenditures (and lower spending levels in 1980s). Neither was there any tendency for achievement scores to increase more rapidly in the states that increased per pupil expenditures most rapidly.

In summary, educational research indicates that the rapid growth in real educational spending during the past two decades has been associated with declining levels of student achievement. Changes in student performance achieved by the states that increased educational expenditures the most were not significantly different from the changes in the states that increased expenditures by substantially smaller amounts. *Given the current structure of our educational system,* there is no reason to believe that higher levels of educational spending will improve student performance.

Our Schools Are Failing Because The Structure Is Flawed

Why are our schools failing? Economic analysis indicates that they are failing for the same reasons that the socialist economies of Eastern Europe and the Soviet Union are failing. When production is regulated by politicians (and bureaucrats) rather than by competition for

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 3

ADJUSTED GRADUATION RATE IN 1982 AND 1987 FOR THE 10 STATES THAT INCREASED REAL EXPENDITURES PER PUPIL THE MOST DURING 1970–88 COMPARED TO THOSE THAT INCREASED EXPENDITURES PER PUPIL THE LEAST

	Adjusted Graduation Rate		
Category	1982	1987	Change
10 States with the Largest Increase in Real Expenditures			
Massachusetts	76.4	76.5	+ 0.1
Connecticut	70.6	80.5	+ 9.9
New Jersey	76.5	77.2	+ 0.7
Nebraska	81.9	86.7	+ 4.8
Wyoming	72.4	89.3	+16.9
Alaska	64.3	66.7	+ 2.4
Delaware	68.2	70.1	+ 1.9
Arizona	63.4	64.4	+ 1.0
Maine	70.1	79.3	+ 9.2
Virginia	73.8	74.0	+ 0.2
Average	71.8	76.5	+ 4.7 (3.4)[a]
10 States with the Smallest Increase in Real Expenditures			
California	60.1	66.1	+ 6.0
Utah	75.0	80.6	+ 5.6
Idaho	74.4	78.8	+ 4.4
Hawaii	74.9	70.8	− 4.1
Iowa	84.1	86.4	+ 2.3
South Dakota	82.7	79.7	− 3.0
Arkansas	73.4	77.5	+ 4.1
Montana	78.7	86.2	+ 7.5
Minnesota	88.2	90.6	+ 2.4
Nevada	64.8	72.1	+ 7.3
Average	75.6	78.9	+ 3.3

[a]Indicates the average when Wyoming is excluded.

SOURCE: U.S. Department of Education, *State Education Performance Chart*, 1989.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 4

SAT (OR ACT) SCORES IN 1982 AND 1988 FOR THE 10 STATES WITH THE LARGEST INCREASE IN REAL PER PUPIL EXPENDITURES DURING 1970–88 COMPARED TO THE SCORES FOR THE 10 STATES WITH THE SMALLEST INCREASES IN EXPENDITURES

	SAT (or ACT) as a Percent of National Average		
Category	1982	1988	Percent Change
10 States with the Largest Increase in Real Expenditures			
Massachusetts[a]	99.4	100.2	+ 0.8
Connecticut[a]	100.3	100.4	+ 0.1
New Jersey	97.3	98.3	+ 1.0
Nebraska	108.2	105.3	− 2.7
Wyoming	104.3	103.7	− 0.6
Alaska	101.6	97.9	− 3.6
Delaware[a]	100.4	99.4	− 1.0
Arizona	101.6	102.7	+ 1.1
Maine[a]	99.7	99.1	− 0.6
Virginia[a]	99.4	99.8	+ 0.4
Average	101.2	100.7	− 0.5
10 States with the Smallest Increase in Real Expenditures			
California[a]	100.7	100.4	− 0.3
Utah	100.0	100.5	+ 0.5
Idaho	102.7	102.7	0.0
Hawaii[a]	96.0	98.2	+ 2.3
Iowa	110.3	108.0	− 2.1
South Dakota	103.8	105.3	+ 1.4
Arkansas	96.2	95.2	− 1.0
Montana	106.0	105.9	− 0.1
Minnesota	109.8	105.9	− 3.6
Nevada	99.5	101.1	+ 1.6
Average	102.5	102.3	− 0.2

[a]Indicates SAT state.

SOURCE: U.S. Department of Education, *State Education Performance Chart*, 1989.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

consumers, producers will spend more time catering to political officials and less time satisfying their customers. Since consumers are unable to shift their business to producers who provide them with "value for the price," suppliers have little incentive to produce efficiently, to innovate, or to figure out better ways of doing things. Simultaneously, since producers are operating in a noncompetitive environment, high costs and a poor-quality product (or service) do not lead to the loss of customers. Inefficient suppliers continue to survive and drain the resources of the economy.

Perhaps a thought experiment will help to explain why the current structure of our educational system does not work. Suppose that the government sought to provide us with a "world-class" restaurant system. The following structure was set up to accomplish this goal. A system of local restaurant districts was established. Taxes were levied to finance the cost of the government-operated restaurants in each district. Citizens were not allowed to patronize public restaurants outside their local district. Of course, snobs were free to choose private restaurants as long as they were willing to pay for the cost of the food twice—once as a taxpayer and again as a customer. If a private restaurant was to survive, it would have to figure out how to attract customers from the tax-financed public restaurants that were giving food away to district residents.

It does not take a genius to figure out that this restaurant system would have major problems. Because the managers of the public restaurants would derive their funds from government, predictably they would spend more time trying to satisfy politicians and less time trying to satisfy customers. The incentive to keep costs low would be weak. After all, higher costs and a lower quality of service would merely highlight the need for additional funding of the public sector restaurants. Since consumers cannot take their funding and go elsewhere, they are in an extremely weak position to discipline the district restaurants that are doing a lousy job. The outcome of this organizational structure is predictable. Compared to the current competitive system, food costs would rise and quality would deteriorate. It would not be long before people would be writing about the "crisis" of America's restaurant system.

Most people recognize that this organizational structure will not work in the restaurant business. Why do we think it will work in education? The structure of the system is flawed. Because public schools do not have to meet the competition of rivals, their costs tend to rise. Lacking freedom of choice, educational consumers are unable to reward producers who supply value at a low cost or to penalize those who do not.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Competition among rival suppliers offering a diversity of products is the foundation of America's prosperity. Producers try harder when they have to compete for consumers. When competition is absent, resources are wasted, costs rise, and quality declines. The provision of education is not an exception to these general propositions.

A Voucher Plan to Improve the Educational Opportunities of Students from Low- and Middle-Income Families

Those who are serious about improving the quality of our schools should seek to promote parental choice and competition among schools. An educational voucher system is the easiest method of accomplishing this objective. The concept of an educational voucher system is really quite simple. Rather than providing aid directly to public schools, the government would provide parents with an educational voucher redeemable at any school, either public or private. Empowered with both purchasing power and choice, parents would choose the best school for their children. In some cases, this would be a private school or a public school outside one's current school district.

Under a voucher plan, schools would compete for students, and parents would be free to choose among schools. The best schools—those providing quality education at an economical cost—would thrive and expand. The demand for outstanding teachers would increase, which would add prestige to the teaching profession and help attract quality people into the profession. Competition among schools would encourage innovative ideas and the discovery of more effective teaching techniques. And perhaps even more important, the lousy schools would either change their ways or be driven from the market.

In the past, two major objections have been raised by the opponents of educational vouchers. First, some have opposed the idea because they believed that high-income families would reap the major benefits from the vouchers, at least initially. After all, those with high incomes are overrepresented among families currently sending their children to private schools. The voucher plan would shift this cost to the taxpayer, perhaps to the detriment of educational funding for students from low- and middle-income families. Second, others have expressed a fear that a voucher plan would increase the racial imbalance among schools.

Overcoming the Objections to Educational Vouchers

A voucher plan directed toward low- and middle-income families would overcome the objections raised against vouchers. Consider

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the following plan. Suppose that low-income families, say those with annual adjusted gross incomes of less than $15,000, were given vouchers of $4,500 per student (approximately the current level of operating expenses per student in the United States) redeemable at any public or private school.[3] Similarly, suppose middle-income families were given vouchers worth $4,500 per student minus $150 for every $1,000 of family income above $15,000. As Table 5 shows, the educational voucher would be worth $3,750 for families with an adjusted gross income of $20,000; $2,250 for those with an income of $30,000; and $750 for those with a family income of $40,000. The value of the educational voucher would be completely phased out for family income levels of $45,000 and above.

This voucher plan would enhance the educational opportunities available to children of low- and middle-income families without providing a windfall gain to high-income families who send their children to private schools. In addition, the plan would actually promote racial and economic balance among schools. Since blacks and other minorities are overrepresented among the low-income families receiving the highest valued vouchers, those groups would also be overrepresented among families shifting to private schools

TABLE 5

VALUE OF THE PROPOSED VOUCHER AS FAMILY INCOME INCREASES

Adjusted Gross Income	Value of Voucher[a]
$0–15,000	$4,500
20,000	3,750
25,000	3,000
30,000	2,250
35,000	1,500
40,000	750
45,000 and above	0

[a]For families with an income of less than $15,000, the value of the proposed voucher is equal to $4,500 per student. For families with incomes between $15,000 and $45,000, the value of the voucher would be equal to $4,500 minus $150 per $1,000 of income above $15,000. Thus, the value of the voucher declines to zero at $45,000 of income.

[3]In the United States, educational policy is a responsibility of the individual states. Since the per pupil expenditures vary by state, the maximum value of the voucher would differ among states. Thus, the dollar figures suggested here are designed merely to illustrate the concept.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(and high-quality public schools outside their district).[4] More generally, the representation of children from families with modest incomes would increase among private and elite public schools.

In essence this plan would make educational choice meaningful for low- and middle-income Americans. Currently, the rich use their purchasing power to choose schools that improve the quality of education available to their children. If given the opportunity, low- and middle-income families will do likewise.

The plan would promote competition among schools—a vital ingredient that is largely absent within the current system. If allowed to compete on something approximating a level playing field, private schools and quality public schools will prosper, expand, and improve the quality of education in the United States.[5]

The voucher plan would also promote diversity and allow a larger number of Americans to choose and receive the type of schooling that they really want. Some parents would choose a highly structured school; others would prefer the open-school concept. Some would select a school that stresses religious values; others would opt for secularism in education. Some would choose a school with a tradi-

[4]One of the strongest proponents of educational vouchers for low-income families is state representative Polly Williams of Wisconsin. Rep. Williams, who was the state chairman of Jesse Jackson's 1984 and 1988 presidential campaigns, recognizes that vouchers would empower low-income blacks and provide them with an alternative to public schools that are failing to meet their needs. Wisconsin recently adopted a bill sponsored by Williams providing educational vouchers for low-income families.

[5]Some people worry that student learning at schools receiving vouchers would be low either because parents make poor choices or because the schools would make things easy to attract students. Administration of standardized basic skills tests (which the state already administers) at several grade levels and publication of the results *by school* would resolve this problem. Publication of test results by school should include not only performance at current grade level, but also the student's prior achievement scores at the previous grade level at which the standardized test was administered. These data would provide parents with information on the *change* in achievement level of students attending specific schools. No doubt, some schools would specialize in the provision of accelerated programs or programs designated for disadvantaged students.

The standardized test results by school should be published in a form that would make it easy for parents to determine (and for schools to communicate) the effectiveness of schools for various groups of students, including both gifted students and those with a *prior* record of low-level performance. The vouchers could be limited to elementary and secondary schools willing to participate in the standardized testing program at three or four grade levels. However, other than publication of average basic skill levels achieved by students, state regulation of schools—including local public schools—should be minimal. Publication of the standardized test results in this form would provide parents with better information concerning the quality of schools than they currently are able to obtain. Armed with this information, parents would have both the incentive and resources needed to make wise choices.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tional college prep program, while others would support schools that provide technical and vocational training. Under a system of parental choice, each of these diverse preferences could be satisfied. In effect, an educational voucher system provides proportional representation; it permits members of each group to vote for and receive their preferred type of education. This characteristic is one of the voucher system's most admirable attributes.

In contrast, the public school curriculum tends toward a homogeneous level that fits only the middle spectrum of students. This homogeneity is a major defect. People are different. There is not one best school, or best automobile, or best television program that fits everybody. Some students will do better in one educational environment, while others will flourish in a very different setting. We do not all choose the same church or civic organizations. Why in the world would anyone think that we would all prefer the same type of school?

How much would this voucher plan cost? Interestingly, it would not add much to current expenditure levels. When children from low-income families shift from public to private schools, the cost of the voucher is offset by lower public-school expenditures resulting from a reduction in the number of students in public schools. When middle-income children shift, the cost of the voucher is less than the savings resulting from lower in-district enrollments in the public schools. For upper middle-income families, the cost of providing the voucher is substantially less than the savings resulting from in-district enrollment reductions. For families with incomes above $15,000, switches to a more preferred school actually reduce public-sector educational expenditures.

Of course, some low- and middle-income families already send their children to private schools. Providing vouchers to these families will increase public-sector expenditures. However, this fact raises an interesting question: Do we really want to force these modest income families to pay for education twice? The current system discriminates against these families. It forces them to pay for education once as a taxpayer and again in the form of a tuition payment—merely because they care deeply about the quality of the school that their child attends. This double payment is an injustice that needs to be corrected. The proposed voucher plan would do so.

It is time to quit protecting the public school's monopoly and to begin allowing private schools to compete for students on a level playing field. Studies indicate that private schools provide their students with a quality education at a fraction of the public-school expenditure levels. In fact, the most comprehensive study in this area, conducted by James Coleman of the University of Chicago

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

found that the verbal and mathematical skills of high school students in private schools (mostly Catholic) *increased* 50 percent more than the skills of similar students in public schools. According to Coleman (1989; also see Coleman and Hoffer 1987), the basic skills of private-school students increased three grade equivalents between the tenth and twelfth grades, compared to the two grade equivalents of similar public school students. Coleman believes that the "community environment" and improved relations between parents and school administrators in private schools explains their greater effectiveness relative to the typical public school.

Should parents desiring to send their children to a school operated by a religious institution be eligible for the vouchers? Absolutely. Why would we want to discriminate against a parent and that parent's child because of their religious preferences? The U.S. Constitution protects the free exercise of religion. Surely the free exercise of religion includes the right to choose a school that reinforces the religious values of one's home. The GI Bill's treatment of educational benefits provides an example of nondiscriminatory treatment across schools. Persons attending college at both secular public and private schools *and* religious-affiliated schools were eligible for educational benefits under the GI Bill. The same principle should be applied here. Given choice, some will choose schools with a religious emphasis (including Protestant, Catholic, Jewish, and Moslem schools). Others will choose completely secular schools. Secularism and religious views would be treated equally. One of the marvelous attributes of a free-choice system is that it would allow both the religious and nonreligious to choose and receive the type of school that they prefer.

Voucher Plans and the Status of Educators

Educational administrators and leaders of teachers' unions have often been at the forefront of those opposed to parental choice and educational vouchers. In some respects their opposition is understandable. Parental choice erodes the monopoly power of the public schools and those who administer them. Competition replaces politics as the regulator of educational quality. Compared to the current system, the demand for administrators and teachers who are unable to deliver a quality product will decline.

However, that is only part of the story—the demand for high-quality teachers and administrators will increase. Competition among schools also means competition for the services of first-rate teachers and innovative administrators who are capable of attracting students. In contrast to the salaries of low-quality educators, the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

salaries of outstanding teachers and administrators would rise under a choice system. Similarly, the respect and prestige of professional educators would be enhanced.

It is also important to recognize that the diversity accompanying choice will also expand the opportunities available to teachers. With choice, it will be easier for teachers to obtain employment at a school with an educational program consistent with their own views. Some teachers will prefer teaching in an open-school environment, while others will feel more comfortable in a highly structured setting. Some teachers will consider only secular schools, while others will be attracted to schools with a religious affiliation. Choice will broaden the options available *to teachers,* as well as to students. Thus, it will improve the fit between the availability of positions and the types of jobs preferred by teachers.

Conclusion

President Bush and the nation's governors recently outlined a set of goals for a "world-class" education system. While goals are fine, we must not lose sight of the difference between "goals" and "achievement." Our current problem is not the absence of goals; rather it is the absence of a sensible plan to achieve the goals.

The current education effort is yielding a scandalously low return because the organizational structure of our educational system is unsound. Without structural change, the experience of recent decades indicates that additional spending is unlikely to improve the quality of education. More money will not solve the problem. Fundamental reform—providing for parental choice, competition, and greater flexibility—is needed. Let's get on with the job before we ruin another generation of young Americans.

References

Bishop, John. "Is the Test Score Decline Responsible for the Productivity Growth Decline." *American Economic Review* 79 (March 1989): 178–97.

Center for Educational Statistics. *Digest of Education Statistics.* Washington, D.C.: U.S. Government Printing Office, annual.

Coleman, James S. "Do Students Learn More in Private Schools than in Public Schools?" *Florida Policy Review* 5 (Summer 1989): 9–14.

Coleman, James S., and Hoffer, Thomas. *Public and Private High Schools: The Impact of Communities.* New York: Basic Books, 1987.

Hanushek, Eric. "The Impact of Differential Expenditures on School Performance." *Educational Researcher* 18 (May 1989): 45–51.

Hood, John. "Education: Money Isn't Everything." *Wall Street Journal,* 9 February 1990.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Lapointe, Archie E.; Mead, Nancy A.; and Phillips, Gary W. *A World of Differences: An International Assessment of Mathematics and Science*, Report no. 19–CAEP–01. Princeton, N.J.: Educational Testing Service, January 1989.

National Education Association. *Estimates of School Statistics*. Washington, D.C., annual.

United Nations Education, Scienific, and Cultural Organization (UNESCO). *Statistical Yearbook 1988*. Paris: UNESCO, 1988.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Labor Law Reform: Lessons from History

Charles W. Baird

Introduction

In 1988 labor union membership in the United States declined to 16.8 percent of the labor force. In the private sector the figure was 13 percent, and in government employment it was 37 percent (BNA, 130 LRR 143). By comparison, the corresponding figures in 1970 were 29.6 percent, 29.1 percent, and 32 percent; in 1980, they were 23.2 percent, 20.6 percent, and 35.1 percent (Troy and Sheflin 1985, pp. 3-15, 3-20). The defeat on July 27, 1989, of the United Auto Workers in the certification election at Nissan Motor Manufacturing Corporation's plant in Smyrna, Tennessee, was a widely publicized example of unionism's loss of market share to nonunion labor.

According to unionists, this decline is due to hostile management that has learned to exploit loopholes in the National Labor Relations Act (NLRA). The Labor Reform Bill of 1977–78 was an unsuccessful attempt by politicians who are sympathetic to unionism to plug those alleged loopholes. But the defeat of that measure has not stopped unionist demands for labor law reform. In his 1989 Labor Day speech, Lane Kirkland, the president of the AFL–CIO, said that he would prefer "no law" to the NLRA as it currently exists (BNA, 132 LRR 13). In November 1989, at its 18th biennial convention in Washington, D.C., the AFL–CIO adopted a resolution calling for a major labor law reform that, like the failed 1977–78 bill, would make it more difficult for employers to compete against unions during certification elections and would significantly increase the penalties imposed on employers found in violation of the law.

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economics at the California State University at Hayward. He thanks Kurt Leube, Sudah Shenoy, Steve Shmanske, Shyam Kamath, and Nan Maxwell, as well as an anonymous referee, for helpful comments on earlier drafts.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

I agree that it is time for the Congress to enact and the President to sign significant legislation for labor law reform. But the sort of labor law reform I have in mind is rather different from that envisioned by Lane Kirkland. What I have in mind is reform that will actually protect, in the words of Section 1 of the NLRA, "the exercise by workers of full freedom of association, self-organization, and designation of representatives of their own choosing." The NLRA never has done this, nor was it ever intended to do so. The AFL–CIO wants changes in labor law that will benefit unions; I want changes in labor law that will benefit workers.

The view that what is good for unions must also be good for workers is based on what I consider to be erroneous, but widespread, beliefs concerning the history of the American union movement. According to the conventional view, employers and courts, with the acquiescence of legislators, systematically abrogated the rights of workers and suppressed the development of labor unions, which were formed to defend those rights. It was not until the second third of the 20th century, with the passage of the Norris-LaGuardia Act (1932) and the Wagner Act (the original NLRA of 1935) that, according to this view, the federal government secured workers' rights by promoting, rather than impeding, unionism.

In this paper I will offer a revisionist view of some major events in the history of American unions. This view supports the assertion of W. H. Hutt (1973, p. 23) that the Norris-LaGuardia and Wagner acts were "economic blunders of the first magnitude. They were worked for and acquiesced to under motivations of almost unparalleled sordidness and cynicism combined with the highest, misguided idealism." The paper also provides a model for legislators to follow in the design of legislation to replace those two infamous acts.

First, I will explain the philosophical and theoretical perspectives from which I interpret labor union history and I will compare them with more conventional views. Second, I will analyze three major 19th-century milestones in union history—the Haymarket affair of May 1886, the Homestead lockout and strike of July 1892, and the Pullman strike and boycott of June-July 1894—from those perspectives. Third, I will reexamine two famous union-related Supreme Court cases of 1921—*Duplex Printing* and *Tri-City*—and will argue that the decisions therein are consistent with my revisionist perspectives and ought to be codified in labor reform legislation in the 1990s. Fourth, I will discuss the main features of the Norris-LaGuardia and Wagner acts, pointing out how they deviate from the concept of "the rule of law" exemplified in the two 1921 decisions and explaining why I think certain provisions of those laws ought to be repealed.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Finally, I will discuss a particularly outrageous proposal for labor law reform made by Paul Weiler of the Harvard Law School, and I will summarize my own recommendations for such reform.

Revisionist and Conventional Perspectives

In history and social science, there is no such thing as a simple fact. All facts, beyond mere physical description, are theory laden. An historian tells stories of human actions and events that are really interpretations of those actions and events, and the interpretations depend, to a large extent, on the writer's philosophical and theoretical perspectives. This varying perspective is surely the case with the history of American trade unionism. While a group of historians may agree that on a specific day, at a specific place, a specific person threw a rock through a specific window, different historians in the group will interpret the action differently.

The philosophical perspective from which I will evaluate the labor market and labor unions in this paper is based on the concept of human rights held by the American Founders, especially James Madison and Thomas Jefferson. They subscribed to John Locke's theory of natural rights (Locke [1689] 1924). In order for *X* to be a legitimate human right in the Lockean sense, *X* must be generalizeable to all humans all the time. If Person A asserts right *X* for himself or herself, *X* must apply to all other people in exactly the same way that it applies to A. If the generalized application of *X* to all people gives rise to a logical contradiction, then *X* is not a legitimate right of A or anyone else.

What, then, are the legitimate job-related rights of all people in the labor market? For example, does any worker have a legitimate right to a job, in the sense of an entitlement to employment in general, or any specific employment, irrespective of his or her own actions, without the consent of those providing the employment? If person A has a right to a job in that sense, there must be some other Person B who has the duty to provide the job to A. But if so, A and B have different job-related rights. An entitlement to receive is different from an obligation to provide. Since the right to a job in this sense is not generalizeable without contradiction, it is not a legitimate right. The only legitimate job-related right that A and B (and everyone else) has is the right to *make offers* to provide or to employ labor on whatever terms they wish. Notwithstanding the legal rights created by the NLRA, no one has a legitimate right either to compel another to accept any specific offer or to restrict the offers that others may make (Baird 1988).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Consider the so-called right to strike. If a strike is simply a collective withholding of labor services in the face of unacceptable terms of trade offered by an employer, and in the absence of a prior agreement that such an action would not be undertaken, then there is a legitimate right to strike. Each individual owns his or her labor services and thus can refuse to accept offers to hire those services. If individuals have the right, so, too, does a group of like-minded individuals. But strikes are actually more than this. They are collective withholdings of labor services coupled with attempts to deny the employer access to replacement workers, other suppliers, and willing customers. They are coercive attempts by individuals who wish to withold their own labor services, to deny, often through violence or threats of violence, the voluntary exchange rights of other people.

The theoretical perspective from which I will evaluate the labor market and labor unions in this paper is based on the complementarity of capital and labor as factors of production and on market process theory. Most economists readily accept the idea that labor and capital are both complementary and substitute factors of production. But, because of the process of entrepreneurial discovery and competition, complementarity between labor and capital is far more significant than substitution. When labor-saving substitutions of capital for labor are made, the displaced labor becomes a profit opportunity for an entrepreneur who notices, or imagines, alternative ways to employ it (Kirzner 1973). In pursuit of such profit opportunities, entrepreneurs assemble new complementary packages of capital, labor, and other inputs. This is, of course, the principal reason why technological unemployment is merely a short-run, disequilibrium phenomenon rather than the long-run, inherent contradiction within capitalism that Marx thought it was.

The significance, for the interpretation of labor union history, of the fact that labor and capital are chiefly complements is most clearly spelled out by W. H. Hutt (1973). Because of that complementarity, the only way that owners of capital can exploit sellers of labor is for the former to "shut in" the latter. Capitalists need labor to work with their capital, and the only way they can obtain it is by entering hiring agreements with those who seek employment. But sellers of labor will not consent to employment with Capitalist *A* if they have better offers of employment with Capitalist *B*. Capitalist *A* shuts in sellers of labor by blocking their access to alternative employments, that is, by creating and sustaining monopsony.

Similarly, providers of labor can exploit providers of capital by shutting the capital in. Labor shuts in capital by making it difficult,

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

dangerous, or illegal for capitalists to withdraw their capital from current employments and redeploy it in alternative employments. The recent union success in pushing the adoption of the Worker's Adjustment and Retraining Notification Act (the 60-day plant closing notification law) of 1987 illustrates the phenomenon.

Nonunion labor of a particular type is a substitute for, not a complement of, union labor of the same type. The only way an owner of an input can exploit the owner of a substitute input is for the former to shut out the latter. An employer will not voluntarily hire union labor if equally capable nonunion labor is available at a lower price. Unionized labor exploits nonunion labor by eliminating the lower-price offers of nonunion labor.

Since the employer is denied access to these alternative wage offers, the employer's capital is incidentally shut in with the unionized labor. However, if the capital is nonspecific, and if there are no legal impediments to the mobility of capital, this shutting in is only temporary. The nonunion labor that is permanently shut out of the unionized employment can find nonunionized employment, some of which is newly created through the redirection of capital out of the unionized sector. Nevertheless, shutting out by unions has permanent negative economic consequences.

Strike threat duress affects the composition of capital. Capitalists try to avoid investments in specific capital that is subject to strike threat. Some profitable investment opportunities are foregone in favor of less profitable, but safer, ones. This misallocation of capital implies a diminished flow of aggregate wages (and other incomes).

According to Hutt (1973, chaps. 8–11), history offers no examples of capitalists shutting in labor except for very brief episodes, and fewer examples of nonunion labor shutting out union labor. But there are several examples of attempts by unionized labor, successful only in the short-run, to shut in capital. Examples of unionized labor shutting out nonunion labor, through such devices as the standard rate, strike threat, and violence, are commonplace.

Lest one think the acts of coercive and violent shutting out are the occasional exception to a peaceful rule, the Wharton School recently published a detailed account of union violence throughout the 20th century up to the early 1980s (Thieblot and Haggard 1983). Even more recently, the Pittston Coal Co. had to spend approximately $20 million to defend itself and nonunion workers against picket-line violence during the nine-month strike by the United Mine Workers in 1989 (*Wall Street Journal*, 2 January 1990, p. 3). Not only were violence and threats of violence common in union history, they are still common today.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

If, as I hope to illustrate with several examples below, labor union history is primarily a story of exploitation of some labor by some other labor, rather than a story of labor's reaction to its exploitation by capitalists, labor unions do not deserve the sympathy of would-be reformers. Labor law reform in the 1990s ought to attempt to protect all workers from coercion from *any* source, including from unions.

The perspective upon which the conventional view of union history is based is the doctrine that employers, especially large corporate employers, have a natural bargaining power advantage over unorganized workers. Because trade unions help to rectify labor's bargaining power disadvantage, in this conventional view, they are able to secure gains for labor at the expense of capital. But a worker's bargaining power with an employer depends on the worker's employment alternatives. The claim that unorganized workers have no bargaining power is simply a claim that such workers typically confront monopsonistic or conspiratorial employers.

There is substantial evidence that monopsony was not a pervasive problem in the 19th century. For example, real wages were on a strong upward trend throughout the 19th century, long before trade unions could have redressed the imbalance of bargaining power. Moreover, employee-initiated job switching was frequent and became increasingly common throughout the century. Large firms typically paid higher real wages than small firms. Employers' bargaining associations that were set up to administer a uniform labor policy were usually established specifically in self-defense against the already formed trade unions (Hutt 1973, chap. 5; Reynolds 1984, chap. 3). But if monopsony was a problem for labor in the 19th century, it was Henry Ford, not Samuel Gompers or Eugene V. Debs, who deserves the credit for ameliorating the situation. The mass-produced, inexpensive automobile increased the area of effective job search for most workers and thus eventually made it impossible for any employer to maintain monopsony power.

In the conventional view, little, if any, attention is paid to the obvious conflict between union and nonunion labor. Labor is treated as a homogeneous class with interests necessarily at odds with capitalists. Nonunion workers are dupes of employers and, according to some writers, traitors to their class. They have no legitimate rights to be put in the balance against the rights of unionists. They are just "scabs," "yellow dogs," or, in Britain and South Africa, "blacklegs." In the words of Jack London, "After God had finished the rattlesnake, the toad, the vampire, He had some awful substance left with which He made the scab" (quoted in Reynolds 1984, p. 31). With the moral

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

merit of nonunion labor a non-issue, the pro-union view of trade union history reduces to oft-repeated fictional stories about the long and bitter battle between labor—the deserving underdog—and capital—the greedy oppressor (Hutt 1973, chap. 3).

Three 19th-Century Milestones

In 1989, Representative Michael McNulty (D-NY) submitted H.R. 2949, "Study of Nationally Significant Places in American Labor History, Authorization," for consideration in the 101st Congress. (As of this writing the fate of the bill is yet to be determined.) There is no doubt that Haymarket Square in Chicago, Illinois; the Carnegie Steel Works in Homestead, Pennsylvania; and the railroad yards of Chicago (the center of the storm during the Pullman strike of 1894) are significant places in American labor history. They will be on McNulty's list. Every conventional labor history text cites these places as scenes of heroic battles of workers against capitalists that, although labor lost each battle, made significant contributions to the process by which justice for working people was eventually achieved. In my view, these famous conflicts are not anything in which unionists can justly take pride. If monuments are to be constructed at these sites, they would more appropriately commemorate the brave nonunion workers who fought bitter battles in defense of individual rights against forced cartelization.

Haymarket

The commonly accepted definitive treatment of the Haymarket affair is Paul Avrich's *The Haymarket Tragedy* (1984). In his preface, Avrich (p. xi) writes:

> The Haymarket affair, a pivotal event in the history of both the anarchist and labor movements, began on May 3, 1886, when the Chicago police fired into a crowd of strikers at the McCormick Reaper Works, killing and wounding several men. The following evening, May 4, the anarchists held a protest meeting near Haymarket Square. . . . The last speaker, Samuel Fielden, was concluding his address when a contingent of policemen marched in and ordered the meeting to be closed. Fielden objected that the gathering was peaceful. . . . The police captain insisted. At that moment a bomb was thrown into the ranks of the police, inflicting serious injury. The officers responded by opening fire on the crowd, killing and wounding a number of civilians, as well as some of their own men. Sixty-seven policemen were hurt, eight of whom afterwards died.

This description is, in a sense, accurate, but it is also terribly misleading. It suggests that the strikers were employees of McCor-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

mick; most were not. It suggests that the police were unprovoked on May 3; they were not. It suggests that the strike at McCormick was the labor issue of chief concern to the May 4 protesters; it was not. It suggests that the police had no grounds for ordering the May 4 protest meeting to disperse; they did.

It is not until chapter 13 that Avrich gives his interpretation of the detailed events on May 3 and May 4. It is clear from how he tells the story that in his view the terse summary of events in the preface is how they properly could be remembered.

Avrich's chief concern is with the anarchist movement and the seven self-proclaimed anarchists who were arrested, tried, convicted, and sentenced to hang as accessories to murder before the fact because of the deaths that emerged from the bomb and riot. In the end, four were executed, one committed suicide, and the sentences of the other two were commuted. These two were later pardoned. The identity of the actual bomb thrower has never been established. In Avrich's view, and in the view of most tellers of the story, the anarchists were unjustly treated.

My chief concern is not with the anarchists or with whether they were justly treated. What intrigues me is how the Haymarket affair could ever be considered a milestone in the history of the American union movement. The chief players, the anarchists, had nothing directly to do with the trade union movement. To be sure, they preached the Marxist doctrine of capitalist exploitation of labor, but they were not unionists. It is simply not true that labor leaders called the May 4 protest meeting. The anarchists were eager to use union-management disputes for their own purposes, but they were decried by major labor union leaders such as Samuel Gompers of the American Federation of Labor and Terence Powderly of the Knights of Labor. And yet today the monument in Chicago's Waldheim Cemetery to the seven anarchists is considered by many as a labor movement shrine.

My interpretation of the events begins on February 16, 1886, when the McCormick Reaper Company was shut down by a lockout that was undertaken in anticipation of a strike that was scheduled to begin the following day. According to the *New York Times* (hereinafter, NYT) 17 February 1886, p. 1), "There has been some trouble between the men and their employers for a week past, mainly over the question of wages. The company has conceded every point but one, and that is that five non-union men at work in the foundry be discharged."

McCormick expressed willingness to end the lockout and pay the union workers wages they had already agreed to accept. All the union had to do was drop its demand that the nonunion men be fired and

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

drop its strike threat. Who was exploiting whom? McCormick was not refusing to recognize the union and employ its members on agreed terms. He was insisting only that he had a right also to hire willing nonunion workers.

On February 25 the first reported acts of violence took place. A foreman on his way to work was stopped by union men and threatened with a gun. Later the same day, an engineer and a group of gas fitters were accosted in an attempt to keep them out of the plant. The engineer went in, and the gas fitters went home (NYT 26 February 1886, p. 1). On February 28, McCormick and a committee of union men agreed that work would be resumed on March 1. A group of other union men who wanted to stay out "intimated very strongly that they will make trouble for those who do attempt to resume labor" (NYT 1 March 1886, p. 1).

On March 1, at first only 150 out of 1,000 men who had expressed a desire to return to work did so. "Great crowds of strikers lined Blue Island Avenue, facing the works with the evident intention of intimidating any men expecting to go to work" (NYT 2 March 1886, p. 4). The police ordered the crowd to disperse, and when it did so, 200 additional men entered the plant to work. McCormick took back all the union men who agreed, and had sufficient courage, to resume work (about one-third of the strikers). By May 1 McCormick had replaced all the other strikers with nonunion men. Thus far, there had been very little violence.

Saturday, May 1, 1886, was designated by the leaders of the eight-hour movement as the start of massive nationwide strikes in favor of legislation declaring an eight-hour work day with 10 hours of pay. The movement was made up of several different groups including socialists, anarchists, and unionists. In Chicago the main participants in the May 1 strike were railroad freight handlers and lumber shovers. The eight-hour strike, just like the earlier McCormick strike, illustrates the exploitation of nonunion labor by union labor. The character of the May 1 activities is illuminated by the following account (NYT 2 May 1886, p. 2) of what happened after a rally of striking freight handlers:

> After the meeting adjourned, the strikers made a tour of the railroad freight depots. Wherever men were found at work, they were induced, either by arguments or threats, to quit. At the Lake Shore station the doors and windows were closed, but a striker with a sharp eye discovered that freight handlers were at work inside. The doors were broken down and a crowd 1,000 strong forced its way into the building. A couple of policemen tried to drive the intruders out, but, of course, could not. Captain Buckley and a squad of men had better luck, and the crowd tumbled out into the street.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The following day, Sunday, passed quietly. Then came Monday, May 3. The day began with parades, rallies, speeches, and picketing. August Spies, one of the anarchists who later would be executed because of the May 4 bomb and riot, was invited by the Central Labor Union (an organization that was attempting to coordinate the general strike) to address a rally sponsored by the Lumber Shover's Union. Most of the strikers at the rally had not ever been McCormick employees. They were strikers, as Avrich says, but most were not McCormick strikers. In chapter 13 Avrich acknowledges that the rally itself had nothing to do with the earlier McCormick strike and that the rally was attended by several thousand demonstrators who had nothing to do with McCormick, but the point is lost because of his choice of focus.

The Lumber Shover's rally was held within three or four blocks of the McCormick plant. Before Spies spoke, a representative of the Central Labor Union, Fritz Schmidt, harangued the audience from the top of a car.

> On to McCormicks and let us run every one of the damned scabs out of the city. It is they who are taking the bread from you, your wives, and your children. On to them, blow up the factory, strike for your freedom, and if the armed murderers of the law interfere, shoot them down as you would the scabs [NYT 4 May 1886, p. 1].

Spies' speech was, according to a 1936 account by historian Henry David, "free from revolutionary propaganda and inflammatory utterances" (David 1936, p. 190). He merely exhorted the demonstrators to maintain class solidarity against employers lest their cause be lost. Before he had finished, the McCormick bell sounded, indicating the end of the working shift. Spies tried to get the demonstrators to ignore the bell, but a large group (the NYT reported thousands) of them broke away, armed themselves with stones, and charged off to the McCormick plant. In David's account:

> They at once attacked the strike-breakers, and pitching in to them with sticks and stones, they drove the "scabs" to the factory for safety. The few police stationed there fired their revolvers in a vain attempt to disperse the crowd, which was smashing the windows of the factory. Some shots were also fired by the strikers in answer to those of the police. A telephone call for aid brought first a patrol-wagon with eleven officers, and then a police detail of almost two hundred men. . . . The police reinforcement, making good use of club and revolver, charged the workers, whose resistance crumbled immediately [David 1936, p. 190].

David's philosophical and theoretical proclivities, like Avrich's, led him to render the story in a way that blurs the distinction between

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

"strikers" and McCormick workers, and between demonstrators and "workers." In keeping with long unionist practice, he directs attention away from the issue of the rights of nonstriking workers by using the epithets "scabs" and "strike-breakers." In my view, attention should be focused on the issue of trespasses of the rioters against McCormick's willing workers and McCormick's property. Therein lies the egregious harm that was done.

Then came May 4. Appalled by what he perceived to be an unjust act of bloody capitalist oppression against innocent workers, August Spies composed a "revenge circular" of which 2,500 copies were printed by the anarchist newspaper, *Arbeiter-Zeitung*. The circular was distributed at labor rallies held during the morning and afternoon of May 4. It read, in part,

> REVENGE! WORKINGMEN! TO ARMS!
>
> Your masters sent out their bloodhounds—the police—they killed six of your brothers at McCormick's this afternoon. They killed the poor wretches, because they, like you, had courage to disobey the supreme will of your bosses. . . .
>
> If you are men, if you are the sons of your grandsires, who have shed their blood to free you, then you will rise in your might Hercules, and destroy the hideous monster that seeks to destroy you.
>
> To arms, we call you, to arms!
>
> YOUR BROTHERS [David 1936, pp. 191–192].

A colleague of Spies, Adolph Fischer, who was to share his fate, composed and printed a handbill announcing the protest meeting that had been scheduled for Haymarket Square on the evening of May 4. It read, in part,

> ATTENTION WORKINGMEN
>
> Great mass meeting tonight at 7 o'clock at the Haymarket. . . . Good speakers will be present to denounce the latest atrocious act of the police, the shooting of our fellow-workmen yesterday afternoon.
>
> WORKINGMEN ARM YOURSELVES AND APPEAR IN FULL FORCE:
>
> The Executive Committee [David 1936, p. 194].

Given the events of the preceding day, and the tone of the circular and handbill, it was not inappropriate for the police to be very concerned about what might occur at the announced meeting. It is not unreasonable to assume that the police would be alert to even the faintest danger of riot and bloodshed and would be eager to intervene before things got out of hand.

The last speaker at the May 4 meeting was Samuel Fielden, another anarchist, one of the two who were eventually pardoned by Illinois

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Governor Altgeld. Among many other less-inflammatory things, he told the crowd,

> You have nothing more to do with the law except to lay hands on it and throttle it until it makes its last kick. . . . Keep your eye upon it, throttle it, kill it, stab it, do everything you can to wound it—to impede its progress. . . .
>
> The socialists are not going to declare war; but I tell you war has been declared on us; and I ask you to get hold of anything that will help to resist the onslaught of the enemy and the usurper. The skirmish lines have met. People have been shot. Men, women, and children have not been spared by the capitalists and the minions of private capital. It has no mercy—so ought you [David 1936, pp. 202–3].

It was a few sentences later when the squad of police under the command of Captain Bonfield approached the speaker and ordered the meeting to end. Then the police were hit by the infamous bomb.

It is true, as Henry David and Paul Avrich argue in their books, that the anarchists were not proven to have done anything more than give inflammatory speeches. But, as Justice Oliver Wendell Holmes said in the famous *Schenck* case (249 U.S. 47 [1919]),"The most stringent protection of free speech would not protect a man in falsely shouting fire in a theatre and causing panic." Whether the anarchists ought to be vindicated or not, there is nothing in these events about which unionists can be justly proud.

Homestead

Homestead, Pennsylvania, is located seven miles east of Pittsburgh on the banks of the Monongahela River. On July 6, 1892, a full-scale battle—complete with dynamite, cannon, fire, and guns—took place there. On one side were locked out and striking workers from the Carnegie Steel Works; on the other were private guards hired by Carnegie to protect plant property and ensure the safety of plant staff members who were still working. Today, a 10-foot-high gray slab monument commemorates the battle. Its inscription reads, "Erected by the members of the Steel Workers Organization Committee Local Unions in memory of the iron and steel workers who were killed in Homestead, Pennsylvania, on July 6, 1892, while striking against the Carnegie Steel Company in defense of their American rights" (Wolff 1965, p. 264).

Lockout: The Story of the Homestead Strike of 1892 by Leon Wolff is widely accepted as an authoritative account of what happened. Wolff, although careful to be even-handed in his coverage, is sympathetic to the union point of view. He agrees with the inscription that

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the strikers were defending their American rights. To the contrary, I would say the strikers were violently denying the American rights of others. Of course, that depends on what one supposes the relevant American rights to be. Let us see what happened.

The story began three years earlier, in July 1889. There was then a strike at the Carnegie Steel Works in Homestead during which the strikers assaulted both people and property. The incumbent sheriff of Allegheny County was unable to control the rioters because he could not find sufficient people who were willing to be sworn as deputies. They were afraid for their lives (NYT 13 July 1889, p. 1). At one point threatening strikers drove a posse of about 100 deputies off company property (Bemis 1894, p. 371). That strike ended in a victory for the strikers. The company agreed to a three-year contract that specified a sliding wage scale on the basis of the price of steel. Wages would rise and fall with the price of steel, but a $25 per ton minimum price set a floor to wages.

This agreement was due to expire on July 1, 1892. By that time the economy was sliding into a recession that was to become a depression, and the price of steel was falling. The passage of the Sherman Silver Purchase Act in 1890 raised widespread fear that the United States would abandon the gold standard. This Act, together with the McKinley Tariff Act of the same year, triggered large capital outflows and contracted the monetary base, setting the economy off into what became the worst depression of the 19th century (Friedman and Schwartz 1963, p. 106).

By June 10 the price of steel was $22.50 per ton (Bemis 1894, p. 376). In May the company offered a new contract that specified a minimum steel price of $22 per ton and that would expire on December 31, 1893, instead of June 30. The Amalgamated Iron and Steel Association, representing 800 skilled steel workers out of a total work force of 3,800, turned down the company's proposal. Only 330 skilled workers were affected by the proposed cuts (Taussig 1893, p. 311). According to Taussig, "Judged by the scale of the market rate of wages for work of similar difficulty elsewhere, some of the men were largely overpaid. Some of the leading workmen received very large earnings indeed,—$6.00, $8.00, even $10.00 per working day" (p. 309).

At first the union insisted that the minimum steel price remain at $25. During negotiations in June, the company raised its offer to $23, and the union lowered its to $24. All other classes of skilled workers such as engineers, blacksmiths, and carpenters had reached agreement with the company on pay scales for the 1892–95 period. The Amalgamated, however, was adamant. It refused to accept anything

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

less than the $24 price, and it insisted that the contract expire in June, not December. Mid-year was the busy period for the firm; the winter months were slack. Henry Clay Frick, Carnegie's man in charge, was equally adamant on the other side.

Frick had set June 24 as a deadline for the Amalgamated to accept his last offer. Failing that, on July 1, the Amalgamated workers would be locked out. In the meantime, with the lessons of the 1889 strike in mind, Frick had a sturdy, nine-foot board fence, topped with barbed wire, constructed around the company property. He also arranged for 300 private Pinkerton guards to be brought in, should the need arise, to protect plant property and other workers. After all, in the 1889 strike the county sheriff proved incapable of doing his job.

When July 1, 1892, came, the promised lockout was undertaken. The Amalgamated workers convinced most other workers to break their existing agreements with the company and to strike in support of the locked-out workers. Company officials alleged that this support was obtained through intimidation, and the Amalgamated alleged that it was a spontaneous act of labor solidarity. In any event, almost all of the 3,800 workers ceased work. The strikers formed an Advisory Committee that was headed by Hugh O'Donnell and that proceeded to usurp the powers of government in Homestead (NYT 2 July 1892, p. 1). Sheriff McCleary, the incumbent police authority in Allegheny County, was unable to deter the strikers. They surrounded the Carnegie works and refused to allow anyone other than company officers to enter for any purpose. From the beginning they made it clear that their chief concern was to deny nonunion men access to the plant. All strangers were stopped and interrogated. If any seemed at all suspicious, they were ordered to leave the town or face physical violence, or worse. Sheriff deputies who had been sent to restore order were driven out of town (NYT 6 July 1892, p. 1). According to Wolff (1965, p. 93):

> Every road leading into the borough was blockaded. No person was allowed to enter without a satisfactory explanation. Newspapermen were issued badges for entry and departure, and those whose reports were derogatory were debadged and hustled out of town. Telegraph communications were set up at union headquarters. Railway depots were surrounded by armed guards. As days passed, the picket line along the waterfront was increased to a thousand men who patrolled the river on both sides, five miles upstream, five miles down. . . . Vigilantes, some of whom were mounted, were placed on the surrounding hills.

In the early morning of July 6, 300 hired Pinkerton guards set out on two towed barges along the Monongahela River for the Carnegie

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

works at Homestead. The plant property bordered the river, and Mr. Frick's fence had been constructed with two large extensions into the river, making a secure docking site on company property. The barges were en route when sighted by Amalgated sentries. As dawn was breaking, alarm bells were sounded in the town. By the time the barges approached their destination, a mob of several thousand people, many of whom were armed, had gathered at the docking site.

Wolff (1965, p. 106) tells us:

> The strikers opened up with a rifle, pistol, and shotgun fusillade which did no damage, except for one bullet which shattered windows in the tugboat's pilothouse. As the three vessels continued on their way, swarms of men followed them by running along the shore, firing from close range.

When the boats reached their destination, the mob crashed through the fence, charged into company property, and told the people aboard the boats they were forbidden to come ashore. Some Pinkerton guards, along with their leader, Captain Heinde, tried to disembark. Wolff (1965, p. 110) paints a vivid picture of what happened next:

> Three strikers ran forward; two grabbed the end of the gangplank while the third deliberately lay down upon it, as if to dare the enemy to cross his body. . . . As Heinde was trying to shove the prone man aside, the latter pulled a revolver and shot him through the thigh. The heavy cartridge knocked him over backward. A torrent of gunfire swept the men on the plank. Heinde was hit again . . . another guard named Klein was killed instantly . . . four of the others were wounded.

Thus far the Pinkertons had not fired a shot, but now they began to return fire. Two strikers were killed and several more were injured. The mob retreated behind protective barricades, and the boats were fired upon. The tug escaped with several wounded Pinkertons aboard and headed for Pittsburgh. The remaining guards aboard the two barges were left behind and fell under siege.

All day long the barges were fired upon. Two cannons were used to try to sink them and drown the passengers. When that failed, there was an attempt to set the barges on fire. At one point dynamite was used in an attempt to destroy the barges and their passengers. In the late afternoon the Pinkertons waved the white flag of surrender. The mob wanted to kill them on the spot, but O'Donnell and other Advisory Committee members got the strikers to agree to accept the surrender. The Pinkertons were assured that they would not be harmed, and they came ashore. They were forced to run a gauntlet and were "inhumanly" beaten. As Taussig (1893, p. 315) recounts, "Beaten, bruised, half dead with hunger, wounds, and fright, they

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

were kept in a large rink, or theatre, until midnight, when they were marched under guard of the Amalgamated Association to the railway station, and thence carried by special train back to Pittsburgh."

The Advisory Committee was now in full command of the town. And it ran the town with a heavy hand reminiscent of Robespierre's Committee of Public Safety almost 100 years before. Monitors and sentries were posted all over town. All movement was policed. People, especially strangers and members of the press, were shadowed and interrogated. The press was censored. Some people were arrested and incarcerated simply because of what they had to say (Burgoyne [1893] 1971, chap. 8). The only thing missing was the guillotine.

Although the strikers had possession and control of the company property, they did no damage. In fact, they even repaired the fence they had torn down earlier. The strikers and their leaders were not socialists or anarchists. They were well aware of the complementarity of capital and labor. They were interested in securing that relationship exclusively for themselves. They wanted to shut in the capital and shut out nonunion labor. In this regard, the workers stated:

> We are naturally interested in the plant, for there is where we make our bread and butter, and, besides it is not our policy to injure the property of the company. . . .
>
> There is one thing certain; even should the authorities take charge of affairs, we are not going to let new men come in under any circumstances. That is what we are going to direct our attention to, whether we or the authorities have charge of the plant [NYT 10 July 1892, p. 5].

The nation's press reported several resolutions of support for the strikers. Other labor unions and politicians had formally adopted these resolutions in which the Pinkertons were condemned as beneath contempt and as an invading force. The strikers were only protecting their homes. The Pinkerton guard was singled out as the "most hated enemy of organized labor." Some state legislatures even passed anti-Pinkerton legislation (NYT 11 July 1892, p. 1). It is easy to understand the loathing that unions had for the Pinkertons. The Pinkertons were the chief force available to employers to bridle the unions' shutting out activities.

On July 10 Governor Pattison activated the National Guard and sent it to Homestead to wrest control of the town and company property from the Amalgamated Advisory Committee. By July 15 enough nonunion men had been hired to begin some operations in preparation for reopening. On July 16 all the striking workers, except the leaders and individuals who were known to have behaved vio-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

lently, were invited back to work on terms equal to Frick's last offer before the strike. Almost none of them did so. The company gradually added to its work force by hiring nonunion workers. On July 27 the plant reopened with about 700 newly hired, nonunion men working. The militia ensured that as many willing workers as could be found would not be molested by the strikers. Through August and September, more and more nonunion men were added to the work force. By September 19 all of the militia except for one regiment had been withdrawn, and the civil government had been restored.

In October, however, a fresh wave of violence broke out against the nonunion workers. It was becoming more and more obvious to the Advisory Committee that they had lost the strike. Some of their own people were quietly returning to work. In a last desperate attempt to stave off the inevitable, a bomb was set off in a boarding house in which many nonunion workers lived. One nonunion worker's own house was burned down. Individual nonunion workers caught alone in town were beaten (NYT 7 October to 4 November 1892).

But it was all for nought. On November 18 a large group of strikers openly went back to work, and on November 20 the Amalgamated officially called off the strike. Thereafter, the firm operated on a nonunion basis. Most of the strikers eventually were rehired, but the Amalgamated was finished at Homestead. Each worker signed his own individual hiring agreement with the firm. The company did not require the workers to agree to refrain from union activity, but the company refused to bargain with any union.

On December 12, 1892, an astonishing story of poisoning of nonunion men broke into the headlines. Patrick Gallagher, one of the cooks responsible for the in-plant feeding of Carnegie nonunion workers beginning in August, charged that he and his fellow cook, Robert Beatty, were hired by Hugh F. Dempsey, an official of the Knights of Labor, to poison the workers' food. There had been a mysterious outbreak of illness and death among the new workers in September and October, but no one suspected foul play. Although the Knights of Labor were not directly involved in the strike, several officials of the union, including Terence Powderly, its head, had voiced support for the Amalgamated. Dempsey, Beatty, and Gallagher were all tried and convicted for the alleged act. They all received prison sentences. No evidence was ever presented to prove that the Amalgamated was involved in the conspiracy (Burgoyne [1893] 1971, chap. 19). In any case, the conspiracy was a case of union solidarity run amok.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

On October 10, 1892, Chief Justice Edward H. Paxson of the Pennsylvania Supreme Court charged the Grand Jury in the trial of the strike leaders for treason against the state. In the end none of the defendants were convicted of that serious charge. But, as the following statement reveals, Justice Paxson understood the true nature of the employment relationship and the legitimate rights and obligations thereof:

> The relation of employer and employee is one of contract merely. Neither party has a right to coerce the other into the making of a contract to which the mind does not assent. The employer cannot compel his employee to work a day longer than he sees fit nor his contract calls for, nor for a wage that is unsatisfactory to him. It follows that the employee cannot compel his employer to give him work or to enter into a contract of hire, much less can he dictate the terms of employment. When the negotiations between the parties came to an end, the contract relations between them ceased. The men had no further demand upon the company, and they had no more interest or claim upon its property. . . . [N]or does it make any difference that a large number were discharged at one time; their aggregate rights rise no higher than their rights as individuals. The mutual right of the parties to contracts in regard to wages . . . is as fixed and clear as any other right which we enjoy under the constitution and laws of this state. It is a right which belongs to every citizen, laborer or capitalist, and is the plain duty of the state to protect them in the enjoyment of it [Burgoyne (1893) 1971, pp. 203–4].

Burgoyne interprets these words of the Chief Justice as evidence that he did not understand the true nature of the dispute. I interpret them to imply exactly the opposite.

Pullman

On May 11, 1894, the production workers at Pullman Palace Car Company in Pullman, Illinois, went on strike. They were protesting wage cuts, which Pullman had unilaterally imposed, as well as the failure of the company to reduce rents in company-owned housing. The depression of the 1890s had emerged full blown in the previous year. Although the leasing and operating divisions of the Pullman Company were still operating profitably, its production facilities were not. The company intentionally continued to manufacture cars at a loss to avoid having to lay off its production workers, but by the spring of 1894 this continued employment practice was becoming untenable at the old wage scales. Even at the new scales, losses would continue (Lindsey 1942, pp. 95–100). Many of the workers lived in housing they rented from the company. These workers

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

requested that Pullman lower the rents at the same time that it lowered wages. Pullman refused.

The American Railway Union (ARU) had been formed in 1893 by Eugene V. Debs and others. The union was organized along industrial, rather than craft, lines. That is, it purported to be an organization of all railroad employees. Several craft unions, called "Brotherhoods," had been in existence for several years and were now competing with the ARU for members. Although many of the striking Pullman workers were ARU members, the ARU did not call the strike. In fact, Debs was initially opposed to the strike because he did not think that the strike could be successful under existing economic circumstances (U.S. Strike Commission Report [1895] 1972, p. XXVII).

The strike at Pullman was peaceful in the sense that no overt violence took place. It was, nevertheless, coercive. The strikers surrounded the works, preventing any entry. Pullman had to lay off the 600 workers who did not strike. A sympathetic teller (Lindsey 1942, p. 123) of the tale put it this way:

> As precaution against violence, three hundred strikers were immediately thrown around the Pullman works for the announced purpose of guarding the property against hoodlums, and this protection was furnished night and day until July 6, when the union was relieved of this duty by the military.

The U.S. Strike Commission ([1895] 1972, p. xxxviii) took a more realistic view of this picketing:

> This guarding of property in strikes is, as a rule, mere pretense. Too often the real object of guards is to prevent newcomers from taking strikers' places, by persuasion, often to be followed, if ineffectual, by intimidation and violence.

Still, no violence or intimidation was reported, and the Commission praised the Pullman picketers for their "conduct and forbearance."

Nothing much happened until June 15. Pullman was not eager to settle the strike because it amounted to a cutoff of an unprofitable undertaking. Pullman refused to enter arbitration. On June 15 the frustrated Pullman strikers appealed for support from the delegates at the ARU convention then under way in Chicago. On June 22 the delegates voted to undertake a boycott of Pullman cars on all railroads unless the Pullman company settled the strike on acceptable terms before June 26. George Pullman, the founder and chief executive officer of the company, refused to relent, and the boycott was undertaken.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The ARU instructed its members to refuse to handle any trains that included Pullman cars. Such trains were to be stopped and the Pullman cars were to be detached before trains were permitted to proceed. If the railroads refused to stop using Pullman cars, they would be struck. On June 26 Debs telegraphed the four railroad Brotherhoods, the Benevolent Order of Switchmen, the American Federation of Labor (AFL), and the United Mine Workers for support. Only the latter responded affirmatively (Lindsey 1942, p. 135). The ARU was essentially on its own.

The General Managers' Association (GMA), an organization of the 24 railroad companies that operated through Chicago, refused to honor the boycott. The GMA's position was that since the ARU had no labor dispute with any of the railroads, and since the railroads had contracts with Pullman to carry Pullman cars, if the railroads honored the boycott they would be in breach of valid contracts.

The GMA had been formed in 1886 by the Chicago railroads to handle problems of mutual concern. It was, in a word, a cartel. But, prior to 1893, the GMA had very little to do with labor matters. It focused on coordination of switching, car service, loading and unloading cars, length and gauge of rails, and so on. (Manning 1960, p. 31). The formation of the ARU in 1893 caught the attention of the GMA, which then began to coordinate labor policies. The first coordinated labor policy was to set the "Chicago scale" for switchmen in 1893. When the boycott began, the GMA coordinated the railroads' response. Under GMA policy, any railroad worker who refused to do his job was to be immediately fired and replaced with a nonunion worker. The GMA set up recruiting offices across the country to find willing nonunion workers. And because of the depression, workers were plentiful. Between June 27 and the end of the boycott (mid-July), the GMA provided 2,500 replacement workers to the Chicago railroads (U.S. Strike Commission Report [1895] 1972, p. 251).

No one disagrees that the boycott and strike were the scene of much violence, much destruction of railroad property, and many injuries to people. In Chicago 12 people were killed (Winston 1901, p. 541). The battle centered in Chicago but it extended from the Mississippi to the Pacific. The unionist position was that the violence was perpetrated by mobs of people who were not strikers, and, in any case, would not have occurred at all if President Cleveland had not sent federal troops to Chicago on July 4.

Violence and property destruction, along with interruption of the U.S. mail and blockades of trains in interstate commerce, occurred from the beginning. On July 2, the federal Circuit Court for the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Northern District of Illinois issued an injunction against Debs and the ARU. They were ordered to stop all violent activity and to refrain from interfering in any way with the normal operations of the railroads (Manning 1960, pp. 54–55). When the injunction was read to strikers and their companions at Blue Island, Illinois, on July 2, the crowd roared its disapproval. "To hell with the Government! To hell with the President! To hell with the Court and injunctions!" (U.S. Strike Commission Report [1895] 1972, p. 362). On July 3, U.S. Marshal J. W. Arnold sent the following telegram from Chicago to Attorney General Richard Olney:

> When the injunction was granted yesterday, a mob of from two to three thousand held possession of a point in the city . . . where they had already ditched a mail train, and prevented the passing of any trains, whether mail or otherwise. I read the injunction writ to this mob and commanded them to disperse. The reading of the writ met with no response except jeers and hoots. Shortly after, the mob threw a number of baggage cars across the track, since when [sic] no mail train has been able to move. I am unable to disperse the mob, clear the tracks, or arrest the men who were engaged in the acts named, and believe that no force less than the regular troops of the United States can procure the passage of the mail-trains, or enforce the orders of the courts. . . . An emergency has arisen for their presence in this city [Cleveland (1904) 1913, pp. 25–26].

The city was relatively calm on July 4 when the first troops arrived, but on July 5, 6, and 7 violence and property destruction were widespread (NYT 6–8 July 1894, p. 1). When it became clear that the federal troops would do whatever they had to do to restore and maintain order, including shooting rioters, the violence diminished. On July 8 it appeared that order had been restored and many trains were running unimpaired. To avoid defeat, Debs requested that the AFL and other unions join the fight in the form of a general strike. On July 12, Samuel Gompers, president of the AFL, formally turned down the request. The strike was lost. Men began returning to work in droves. On July 18 the federal troops were removed. The ARU boycott and railroad strike were over. The Pullman strike itself was ended on August 2. "By August 24 there were twenty-three hundred employees at work in Pullman, of whom five hundred and fifty were new men" (Lindsey 1942, p. 270).

Conventional accounts of the "Pullman strike," as the entire incident became known, focus on three issues: (1) the dispute between President Cleveland and Governor Altgeld as to whether Cleveland acted within his authority when he sent in the troops without first getting clearance from the governor, (2) the legitimacy of the injunction issued against the ARU, and (3) whether the ARU was responsi-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ble for the violence. The U.S. Supreme Court (as we will see below) decided the first two issues at the time by supporting President Cleveland's actions and upholding the penalties assessed against Debs for violating the July 2 injunction. Today, of course, historians who are sympathetic to unionism are critical of the Supreme Court's actions. Post–New Deal Supreme Courts would, most likely, have decided differently.

The consensus among historians is that very little of the violence against railroad property can be attributed to strikers and the ARU. The U.S. Strike Commission ([1895] 1972) has assigned the blame for most property destruction and rioting to unruly mobs made up mostly of "hoodlums, women, a low class of foreigners, and recruits from the criminal classes. Few strikers were recognized in these mobs, which were without leadership, and seemed simply bent upon plunder and destruction" (pp. XLV–XLVI).

The Commission did ascribe some property-related violence to the strikers. The strikers were guilty of "spiking and misplacing of switches, removing rails, crippling of interlocking systems, the detaching, side tracking and derailing of cars and engines, placing of coupling pins in engine machinery, blockading tracks with cars, and attempts to detach and run in mail cars" (p. XLV). But they were not responsible for the large-scale burning and destruction that took place.

There remains an essential question: How did the strikers deal with nonunion replacements and dissenting unionists? It is in the interests of strikers to refrain from too much destruction of capital, but it is also in their interest to shut out substitute labor. This strike, like Haymarket and Homestead, was primarily a contest between some labor and other labor. The union was not defending the rights of labor, it was asserting and defending privileges of some union labor against some other union labor and nonunion labor.

The U.S. Strike Commission ([1895] 1972) itself had very little to say about the battle between some labor and other labor. It simply acknowledges that "considerable threatening and intimidation of those taking strikers' places were committed or instigated by strikers" (p. XLV). Appendix A of the Commission's Report (in the recorded statements of individuals testifying on behalf of the railroads) gives details of the struggle between union and nonunion labor. There one reads stories of whipping, beating, white-washing (covering the victim with white-wash), and intimidating people. A. P. Winston (1901, p. 546) explained the strikers' attitude toward nonunion labor with these words: "The trade union movement partakes essentially of the revolutionary character when 'scabs' are to

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

be dealt with. This class of enemies strikes at the very existence of the union and the union, therefore, looks upon them as outlaws."

The attitude of Debs toward nonunion workers was not in doubt. In one telegram Debs warned a union official in Indiana that "the Baltimore and Ohio officials are . . . looking for scabs in the eastern cities. By all means have them shut off." Another telegram bearing Debs' name, but which he later denied he sent, says "all men continuing in the employ of the railroad departments . . . will be forever branded as scabs, and will be treated as such" (von Holst 1894, p. 507).

The attacks against non-ARU labor were sufficiently numerous and serious that the July 2 injunction specifically singled them out. Items 7 and 8 in the injunction enjoined:

> (7) Compelling or inducing, or attempting to compel or induce by threats, intimidation, persuasion, force, or violence, any of the employees of any of said railroads to refuse or fail to perform any of their duties as employees. . . .
>
> (8) Compelling or inducing, or attempting to compel or induce, by threats, intimidation, force, or violence, any of the employees of any of said railroads, who are employed by such railroads . . . to leave the service of such railroad [Manning 1960, pp. 54–55].

The enjoining of "persuasion" in item 7 is especially interesting. While many people would agree that threats, intimidation, force, and violence against non-ARU labor were properly enjoined, many would also say that there is nothing wrong with some workers attempting to persuade other workers to join a strike. Since threats and violence were specifically enjoined, the judge's objection to persuasion could not merely be that "persuasion" often takes the form of threats and violence. His concern could well have been that even friendly persuasion to join a strike amounts to encouraging the breach of valid contracts. If nonstriking workers were under written or verbal contract to continue working, strikers attempting to get the nonstrikers to join the walkout were guilty of subornation. Subornation of contracts was also the claim of the railroads when they refused to boycott Pullman cars: Extant contracts between Pullman and the railroads precluded the boycott. Only someone who believes that the rights of ARU labor overrode the rights of all the other people involved could justify such subornation. As we will see below, the use of "peaceful persuasion" by unionists was taken up by the Supreme Court in 1921.

Union sympathizers should not forget that other labor unions, especially the railroad Brotherhoods, did not support the ARU. For exam-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ple, Winston (1901, p. 557) reports that on July 3, H. E. Wilkinson, leader of the Brotherhood of Trainmen, made a public statement regarding the ARU strike in which he claimed that the strike was called by delegates

> who did not represent one thirtieth of the employees in train service in the United States, but every man, woman, and child employed in any capacity on a railway is expected to bow to this imperious command regardless of any rights of their own, obligations to other organizations, or contracts with their employers.

Debs condemned the Brotherhoods as "active allies of the railroads in the great strike" (Winston 1901, p. 557), and today Debs is accepted as the spokesman for "labor" in the strike. Dreams of class solidarity notwithstanding, there was no labor point of view. Rather, competing points of view were held by many different sellers of labor services.

In *In Re Debs* (158 U.S. 564 [1895]), the U.S. Supreme Court upheld the penalties imposed by the lower court on Debs on the grounds of the interstate commerce clause of the Constitution. Thus began, according to unionist history, a long period of "government by injunction," a period of oppression of unions through anti-strike injunctions. The period was brought to a just end only by the Norris-LaGuardia Act of 1932.

Two Key Supreme Court Decisions

There were many anti-union federal and state court decisions from 1895 through 1921. But all of the decisions were based on unions' specific actions that, under the ordinary rule of law, were deemed impermissible. None of the decisions were based on the formation of unions per se or on actions of unions that did not trespass against the voluntary exchange rights of others. Freedom of association was defended for all. That freedom of association includes the right of any individual, as a matter of individual choice, not majority vote, to decline to associate with private groups of which he or she disapproves or with which he or she shares no values. The right to form, join, and participate in labor unions and the right to refrain from doing so received genuine equal protection under the law. Two U.S. Supreme Court decisions—*Duplex Printing Press Co. v. Deering* (254 U.S. 443 [1921]) and *American Steel Foundries v. Tri-City Central Trades Council* (66 L.ed. 189 [1921])—illustrate the thinking of the Court during this period particularly well.

In 1908 the Supreme Court applied the Sherman Antitrust Act against union actions that the Court held were in restraint of trade. The AFL was so incensed at the Court's decision that it mounted an

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

effective special-interest campaign aimed principally at Democratic politicians. The intent of the campaign was to get the AFL exempted from antitrust laws (Gould 1987, pp. 15–16). In 1912 the Democratic Party officially promised to enact such legislation if its candidate for president, Woodrow Wilson, was elected. He was, and in 1914 the Clayton Antitrust Act became law. Section 20 of the law states

> that no restraining order or injunction shall be granted by any court of the United States . . . in any case between an employer and employees, or between employers and employees, or between employees, or between persons employed and persons seeking employment, involving, or growing out of, a dispute concerning terms or conditions of employment unless necessary to prevent irreparable injury to property, or to a property right, of the party making the application. . . .
>
> And no such restraining order or injunction shall prohibit any person or persons, whether singly or in concert, from terminating any relation of employment, or from ceasing to perform any work or labor, or from recommending, advising, or persuading others by peaceful means so to do.

Samuel Gompers, president of the AFL, hailed the Clayton Act as "Labor's Magna Carta," but his joy was short-lived.

In the *Duplex* case, the Supreme Court upheld an anti-union injunction on the grounds that the Clayton Act protected unions only when they were "lawfully carrying out their legitimate objectives" (Gould 1987, p. 18). The union called a strike because the company operated on an open-shop basis. That is, the company refused to discriminate against any workers whether they were unionists or not. The union wanted the employer to discriminate against nonunion workers. Duplex continued to operate successfully in spite of the strike, so the union turned to secondary activity to try to get its way. The union threatened commercial customers of Duplex with sympathetic strikes unless they boycotted the company. The union also ordered all repair shops to stop doing business with Duplex on pain of sympathetic strikes. Union employees of the repair shops were threatened with loss of union membership if they did any work on Duplex presses. The Supreme Court held such secondary activity to be actionable under the antitrust laws because the Clayton Act protected only union strike action directed against the specific employer involved in the primary strike. It was not a "legitimate objective" for the union to impair the voluntary exchange rights of third parties. There was no question of the Court intervening in attempts of a group of employees of a firm attempting peacefully to obtain better terms of trade from their employer. The Court was

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

exercising its responsibility to prevent aggression by strangers attempting to spread their empire.

In the *Tri-City* case, the Supreme Court spelled out detailed rules to be applied during labor disputes. I think these rules should be incorporated in any reform of American labor law in the 1990s. In November 1913 the company closed its doors because of bad business conditions, and it laid off about 1,600 men. Some of these men were members of unions affiliated with the Tri-City Council, but the company had always operated on a nondiscriminatory open-shop basis. In April 1914 the company reopened and resumed operations on a reduced scale. Three hundred and fifty former employees, some of them union members, were hired back. The pay scales paid to all workers were reduced by two cents to ten cents an hour compared to the scales paid before the shutdown. The Tri-City Council, made up almost entirely of people who had never been employed by the company, called a strike in protest of the wage cut. Only two of the rehired employees joined the strike. The remaining 348 workers wanted to continue to work, but the Tri-City Council set up large numbers of stranger pickets at every entrance to the firm. Between April 22 and May 18 these stranger pickets repeatedly assaulted and beat workers who tried to continue working. Several employees ended up sleeping in the plant because they were afraid of what might happen if the pickets caught them out in the open. On May 18 a restraining order was issued by a federal court, and the violence stopped. The union argued that Section 20 of the Clayton Act precluded the restraining order against the picketing. The Supreme Court disagreed.

The Court reaffirmed its ruling in *Duplex* that the Clayton Act protected unions when they were lawfully pursuing their legitimate objectives and went on to say that the law protected only "peaceable persuasion by employees, discharged or expectant" (at 197). Strangers had no standing at all to participate in the strike. Strangers could be enjoined even when using peaceable persuasion of workers who wished to work. All violent acts and threats by employees could be enjoined. Only nonthreatening discussion between nonstrikers and striking employees was immune to injunction. Since even peaceful picketing in large numbers is inherently threatening, the Court held, only one picket per entrance could be allowed. Moreover, the Court declared that the right to carry on a business is a property right. An employer is, therefore, entitled to the same protection from people who trespass against that right as the employer is from trespass against any other property right. The *Tri-City* decision is worth quoting at some length:

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

> The irreparable injury to property or to a property right, in the first paragraph of Section 20, includes injury to the business of an employer, and . . . the second paragraph applies only in cases growing out of a dispute concerning terms and conditions of employment between an employer and employee[s] . . . and not to such dispute between an employer and persons who are neither ex-employees nor seeking employment [at 196].
>
> The object and problem of Congress in Section 20, and indeed, of courts of equity before its enactment, was to reconcile the rights of the employer in his business and in the access of his employees to his place of business and egress therefrom without intimidation or obstruction, on the one hand, and the right of the employees, recent or expectant, to use peaceable and lawful means to induce present employees and would-be employees to join their ranks, on the other . . . [at 197].
>
> How far may men go in persuasion and communication, and still not violate the right of those whom they would influence? In going to and from work, men have a right to as free a passage without obstruction as the streets afford, consistent with the right of others to enjoy the same privilege. We are a social people, and the accosting by one of another in an inoffensive way, and an offer by one to communicate and discuss information with a view toward influencing the other's action, are not regarded as aggression or a violation of that other's rights. If, however, the offer is declined, as it may rightfully be, then persistence, importunity, following and dogging, become unjustifiable annoyance and obstruction which is likely soon to savor of intimidation. From all of this the person sought to be influended has a right to be free, and his employer has a right to have him free [at 197].
>
> Tri-City Central Trades Council and the other defendants, being neither employees nor strikers, were intruders into the controversy, and were engaged without excuse in an unlawful conspiracy to injure the American Foundries by enticing its employees, and, therefore, should be enjoined [at 199].

The Court said that its *Tri-City* ruling did not create any new law or new perspective on existing law. It was "merely declaratory of what was the best practice always" (at 197). The basic idea is that every person, whether a union worker, a nonunion worker, an employer, a customer, or a supplier, has equal rights under the law. Each person has a right to make voluntary exchange offers to anyone else, but no person has a right to coerce another to accept his or her offers. Voluntary exchange requires mutual consent. Each person must be able to say "no" to an unacceptable offer from another and to walk away without being molested. In the absence of an explicit agreement to the contrary, the employment relationship gives the employee a right to seek, in nonviolent ways, better terms from the

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

employer. An individual employee may, acting alone or in concert with other employees, withhold labor services in the face of unacceptable terms of employment offered by the employer. If this is what a strike is, then there is a right to strike. But no person has a right to prevent willing workers from working, willing customers from buying, and willing suppliers from delivering during a strike. Moreover, perfect strangers have no standing with the employer. The employer has a right to refuse to deal with them and to prevent them from trespassing against his property. In 1921 the Court thought all this was common sense. I still think it is.

The Triumph of Interests over Rights

The evolution of the majority opinion of the Supreme Court away from support of individual voluntary exchange rights and toward the approval of governmental and union abrogation of those rights has been well reported and analyzed (e.g., Siegan 1987 and Dickman 1987), but space does not permit a discussion of that sad story here. Suffice it to say that the Norris-LaGuardia Act and the NLRA constitute significant departures from the ordinary rule of law as it was understood by the U.S. Supreme Court before the mid-1930s. Even the present Court considers these acts to be consistent with the U.S. Constitution, notwithstanding the fact that they constitute a body of legislated rules that grant special privileges to some at the expense of many. The unique characteristics of labor law, set aside from ordinary principles of fair play, have been aptly called "the apartheid of labor law" (Vieira 1986, p. 35).

Norris-LaGuardia Act

The Norris-LaGuardia Act did four things: It gave legal standing to unions in labor disputes even when they have no members who are employees of the firms involved in the disputes (Section 13), it banned all federal injunctions in all labor disputes under all circumstances (Sections 1, 7–12), it granted labor unions blanket immunity to all antitrust laws (Sections 4–5), and it made union-free contracts between employees and employers unenforceable in federal courts.

As subsequent U.S. Supreme Court cases—e.g., *U.S. v. Hutcheson* (321 U.S. 219 [1941]) and *Apex Hosiery v. Leader* (310 U.S. 409 [1940])—revealed, the Norris-LaGuardia Act effectively insulated labor unions from any, including criminal, prosecution for activities during labor disputes. Unions became unique organizations with very special privileges that no other organizations could ever hope to receive. And according to my reading of history, there was absolutely no justification for the granting of such special privileges.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

For example, the concern over the misuse of injunctions in labor disputes was unfounded. According to Sylvester Petro (1978), in the period from 1880 to 1932 federal injunctions were issued in less than 1 percent of all work stoppages. State injunctions were issued in less than 2 percent of all work stoppages. Moreover, no peaceful primary strike was *ever* enjoined during that period. All court interventions in labor disputes were to enjoin acts of violence including physical harm to persons and property. The few injunctions that were issued were intended to ensure that the same rules of just conduct apply to all parties in labor disputes.

The issue of union-free (or, as unionists like to call them, "yellow dog") contracts raises significant questions. These contracts were agreements between employers and employees that both groups would avoid involvement with outsider unions. According to Reynolds (1984, chap. 5), such contracts were often initiated by *employees* who were seeking protection from the attempts of hostile strangers to interfere with their rights to sell their own labor services on terms they found satisfactory. Of course, union-free contracts were also often initiated by employers who were seeking to avoid having to deal with organizations of hostile strangers.

No matter who initiates such contracts, under the principle of freedom of contract, and in the absence of fraud, there is nothing unjust about them. Under freedom of contract, employers may offer whatever terms and conditions of employment they wish to prospective employees. The prospective employee may accept or reject the offered terms. Similarly, employees may ask any terms and conditions of employment they wish from prospective employers. For example, a prospective employee may ask that the employer hire only union members or that the employer refrain from hiring any union members. The prospective employer may accept or reject the offered terms. No worker has a right to be employed by any specific employer, and no employer has a right to the labor services of any specific employee. The only right anyone has is to make offers to others and to see if the offers are accepted.

But this view of the question is based on the idea that the same rules ought to apply to everyone irrespective of station or status. According to Section 2 of the Norris-LaGuardia Act,

> the individual unorganized worker is commonly helpless to exercise actual liberty of contract and to protect his freedom of labor, and thereby to obtain acceptable terms and conditions of employment.

Therefore, special rules ought to apply. What counts is the result, not the process. Section 2 implies that one does not have freedom of

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

contract unless one is "free" to get the terms one wants to get. But, then, how about workers who want to be free of union involvement? Does not the impairment of union-free contracts make them less free because they cannot get the terms they want? Such conflicts of freedom are inevitable when freedom is redefined as it is in the Norris-LaGuardia Act.

This idea, that workers without unions will inherently have a disadvantage in bargaining power relative to employers, is the basis for most individuals' support of unionism and is picked up again in the Wagner Act. But that disadvantage is a hoary myth. A worker's bargaining power depends on the worker's alternatives. If a worker either works for Employer *A* or does not work (i.e., if Employer *A* is a monopsonist), the worker has little bargaining power. If the worker has several employment alternatives, he has strong bargaining power. There may have been instances of monopsony or oligopsony in the 19th century, but, as we saw above, they were short-lived. Monopsony has not been a significant factor in the American labor market since the introduction and widespread use of the automobile.

The National Labor Relations Act

Section 1 of the National Labor Relations Act (NLRA) begins with an often-repeated fiction that amounts to the identification of employers as an enemy class:

> The denial by employers of the right of employees to organize and the refusal by employers to accept the procedure of collective bargaining lead to strikes and other forms of industrial strife. . . .

In fact, employers rarely denied workers the right to organize before 1935. Most often, they simply refused to recognize outsider unions as exclusive bargaining agents (Baird 1984, chap. 2). Many employers recognized outsider unions on a proportional representation basis even before Section 7(a) of the National Industrial Recovery Act required them to do so. In cases involving union-free hiring contracts, employers refused to deal with any outsider unions, but that refusal was often at the behest of workers who wished to maintain a union-free environment. The industrial strife that was alluded to occurred mainly during recognition strikes as outsider unions attempted to extend their empire over unwilling workers.

Section 1 of the NLRA ends with Orwellian doublespeak in order to justify creating adversarial relations between worker majorities and worker minorities. The NLRA asserts that its intent is to protect "the exercise by workers of full freedom of association, self-organization, and designation of representatives of their own choosing" for

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

collective bargaining purposes. But Section 9(a) of the NLRA destroys freedom of association by making it possible for outsider unions to compel unwilling individual workers to accept their representation services, and Section 8(a)3 compels workers to join outside unions and to support them financially. As such, the NLRA abolishes self-organization and free choice of representatives by forcing all individual workers to submit to the will of the majority, notwithstanding the long-standing common law principle that the sale of one's labor services is in the private, not governmental, sphere of human action.

When a worker decides whether to employ an attorney to represent him or her in a court action, the selection is not made by majority vote, it is made by individual choice. In contrast, under the apartheid of labor law constructed by Section 9(a)'s principle of exclusive representation, when a worker decides whether to be represented by a union in the sale of his or her labor services, individual free choice is overridden by majority vote. Since unions are private organizations, not governments, mandatory submission by an individual to the will of a majority is unjustifiable. It is the triumph of interests over rights.

When the bill that became the NLRA was debated in Congress, Senator Miller Tydings offered an amendment that would have outlawed employee coercion from *any* source, whether from employers or from unions. Senator Tydings said:

> My theory is that there may be men who do not want to join the union who will be coerced under the Senator's bill into joining a union where really they are happy and contented as they are and do not want to be forced to join a union [Dickman 1987, p. 268].

But Senator Wagner assured the Senate that representation elections precluded coercion from unions even in the case of the closed shop (where a worker must be a union member to be hired), and the Tydings amendment was defeated.

The intent of the NLRA was not to ensure employee free choice at all; the intent was to make it possible for unions to draft members and to force employers to deal with hostile outsiders (Petro 1957, chap. 8). The law promotes an adversarial relationship not only between unions and management but also between workers who are union sympathizers and workers who wish to remain union free. Of course, before the 1930s there was strife between union and nonunion workers. But the law was on the side of individual free choice. The law gave special privilege to neither group. When Section 1 states, "experience has proven that protection by law of the right of

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

employees to organize and bargain collectively . . . remov[es] certain recognized sources of industrial strife" and promises that the NLRA will achieve that end, it is merely promising the peace that comes from surrender. And it is workers who wish to remain union free who are forced to surrender.

Exclusive representation gives rise to another injustice in American unionism—the incessant effort of unions to force all the workers they represent to either join them or pay dues to them as a condition for continuing employment. Section 7 of the NLRA states that all employees have a right to affiliate with unions "of their own choosing" and to "refrain from any or all such activities *except to the extent that such right may be affected by an agreement requiring membership in a labor organization as a condition of employment as authorized in section 8(a)3*" (emphasis added).

Section 8(a)3 spells out the details of such compulsory unionism arrangements. The unionist argument in favor of such union security compulsion is that since Section 9(a) compels a certified bargaining agent to represent all the workers in a bargaining unit, all such workers should be forced to pay their fair share of the cost of union services. Otherwise, some workers would get the benefits of union services for free—they would be "free riders." The obvious reply is that Section 9(a) should be repealed. If there were no exclusive representation, there could be no free riders. Unions would bargain for their members only.

The Norris-LaGuardia Act made union-free (yellow dog) contracts between employers and employees unenforceable in federal courts. The NLRA goes even further and outlaws such agreements. In contrast, the NLRA permits forced unionism agreements between employers and unions at the expense of dissenting workers. And, hypocritically, it does so in the name of employee's free choice. Consistency requires that both union-free and forced unionism agreements be permitted, or both be outlawed. In the absence of an exclusive representation, where an individual employee's free choice of whether to affiliate with a union was guaranteed with no exceptions, I would prefer to see both types of arrangements permitted. If employees and employers could self-select the environments within which they work together, labor-management cooperation would be much easier to achieve and entrepreneurial discovery of superior modes of labor-management relations would be facilitated (Baird 1987).

Conclusion

Paul Weiler, of Harvard Law School, a prominent advocate of labor law reform, argues that certification elections ought to be "instant"

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(Weiler 1983, pp. 1811–16). That is, the employer ought to have no opportunity to make a case in favor of a nonunion vote. Weiler's reasoning is astonishing. He says that an employer's participation in a representation election campaign is as inappropriate as a foreign government's participation in an American election. "Only the employees' interests are supposed to count in the certification decision," argues Weiler; thus, "the employer can claim no positive right to influence the employees' vote" (p. 1814).

Weiler's view, apparently, is that a job is a property right of the employee who holds it. Employers who express their own views regarding the collectivization of bargaining trespass against that property right. The more common and, I think, more defensible position is that a job is a contractual relationship between an employee and an employer. Both have a legitimate stake in the outcome of the collectivization issue. Both are properly free to participate in the process by which that outcome is determined.

Weiler tells us that, under the NLRA, labor has "group rights" that supersede the rights of individuals (p. 1788). Ironically, this is exactly how P. W. Botha, the erstwhile president of South Africa, defends apartheid. Each racial group is assigned group rights by the South African constitution. Individual rights, such as freedom of speech and contract, are overridden by these group rights.

Labor law reform ought to be directed at ensuring equal treatment under the law for independent workers, unionized workers, and employers. It should not favor union workers over everyone else. The certification election reform that ought to be adopted is to abolish certification elections altogether. Let individual workers choose on an individual, not a majority vote, basis whether to be represented by a union. Let unions represent their voluntary members and no one else.

At the very least, Section 9 of the NLRA ought to be amended to provide for regularly scheduled, periodic recertification elections of exclusive bargaining agents. As it is now, once a union is certified, it is presumed to continue to have majority support forever unless the contrary is proven by a decertification election. The latter is held only if 30 percent or more of the workers who are represented by an exclusive bargaining agent sign cards requesting that it be held. It is as if once President Bush is elected, he remains president indefinitely unless enough citizens sign a petition requesting an election to determine whether to depose him.

Even with the special privileges they enjoy under present labor law, unions are losing their market share to nonunion labor. If labor law were reformed to eliminate all those special privileges—most

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

importantly, if individual workers were free to decide on an individual basis whether to be represented by a union, and if strikers could withhold only their own labor services and be effectively prohibited from interfering in any way with the voluntary exchange rights of nonstriking employees, suppliers, and customers of struck firms—then unions would lose even more of their market share. But the loss of unions' market share is no excuse to propose, like Weiler, even more drastic special privileges to attempt to reverse the unions' falling market share. After all, why should unions be considered so worthwhile that they should be given special privileges? Their history does not support such a sentiment.

References

Avrich, Paul. *The Haymarket Tragedy.* Princeton: Princeton University Press, 1984.

Baird, Charles W. *Opportunity or Privilege: Labor Legislation in America.* Rutgers State University: Transaction Books, 1984.

Baird, Charles W. "Labor Law and Labor-Management Cooperation: Two Incompatible Views." *Cato Journal* 6 (Winter 1987): 933–50.

Baird, Charles W. "The Varieties of 'Right to Work': An Essay in Honor of W. H. Hutt." *Managerial and Decision Economics,* Special (Memorial) Issue (Winter 1988): 33–43.

Bemis, Edward W. "The Homestead Strike." *Journal of Political Economy* 2 (June 1894): 369–96.

Bureau of National Affairs (BNA). *Labor Relations Reporter* (LRR), Vol. 130, 6 February 1989.

Bureau of National Affairs (BNA). *Labor Relations Reporter,* Vol. 132, 4 September 1989.

Burgoyne, Arthur G. *Homestead: A Complete History of the Struggle.* 1893. Reprint. New York: Augustus M. Kelley, Publishers, 1971.

Cleveland, Grover. *The Government in the Chicago Strike of 1894.* 1904. Reprint. Princeton: Princeton University Press, 1913 (the text of a speech delivered at Princeton University).

David, Henry. *The History of the Haymarket Affair.* New York: Russell & Russell, 1936.

Dickman, Howard. *Industrial Democracy in America.* La Salle, Ill.: Open Court, 1987.

Friedman, Milton, and Schwartz, Anna J. *A Monetary History of the United States, 1867–1960.* Princeton: Princeton University Press, 1963.

Gould, William B. *A Primer on American Labor Law,* 2nd ed. Cambridge: MIT Press, 1987.

von Holst, W. "Are We Awakened?" *Journal of Political Economy,* 2 (September 1894): 485–516.

Hutt, William H. *The Strike Threat System.* New Rochelle: Arlington House, 1973.

Kirzner, Israel M. *Competition and Entrepreneurship,* Chicago: University of Chicago Press, 1973.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Lindsey, Almont. *The Pullman Strike.* Chicago: University of Chicago Press, Phoenix Books, 1942.
Locke, John. *On Civil Government.* 1689. Reprint. London: Dent, 1924.
Manning, Thomas G. *The Chicago Strike of 1894.* New York: Henry Holt and Company, 1960.
Petro Sylvester. *The Labor Policy of the Free Society.* New York: Ronald Press, 1957.
Petro, Sylvester. "Injunctions and Labor Disputes: 1880–1932." *Wake Forest Law Review* 14 (June 1978): 341–576.
Reynolds, Morgan O. *Power and Privilege: Labor Unions in America.* New York: Universe Books, 1984.
Siegan, Bernard H. *The Supreme Court's Constitution.* New Brunswick, N.J.: Transaction Books, 1987.
Taussig, F. W. "The Homestead Strike." *Economic Journal* 3 (June 1893): 307–18.
Thieblot, Armand J., and Haggard, Thomas R. *Union Violence: The Record and the Response by Courts, Legislatures, and the NLRB,* Philadelphia: The Wharton School, University of Pennsylvania, 1983.
Troy, Leo, and Sheflin, Neil. *Union Sourcebook,* 1st ed. West Orange, N.J.: Industrial Relations Data and Information Services, 1985.
U.S. Strike Commission. *Report on the Chicago Strike of June-July 1894.* 1895. Reprint. Clifton, N.J.: Augustus M. Kelley, 1972.
Vieira, Edwin, Jr. "From the Oracles of the Temple of Janus: *Chicago Teachers Union v. Hudson.*" *Government Union Review* 7 (Summer 1986): 1–37.
Weiler, Paul. "Promises to Keep: Securing Workers' Rights to Self-Organization Under the NLRA." *Harvard Law Review* 96 (June 1983): 1769–1827.
Winston, A. P. "The Significance of the Pullman Strike." *Journal of Political Economy* 9 (September 1901): 540–61.
Wolff, Leon. *Lockout: the Story of the Homestead Strike of 1892.* New York: Harper & Row, 1965.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Mobility Barriers and Public Policy

Martin Ricketts

Introduction

This paper discusses the relationship between the "barriers" concept in industrial organization and in the area of geographical mobility. It is argued that Demsetz's (1982) critique of the barriers literature has a wider relevance that is not generally understood. Two traditions now exist side by side. In the first, a barrier is a departure from an imaginary and frictionless situation (e.g., Lott 1986). In the second, a barrier is a departure from a desirable and attainable situation of optimal friction (e.g., Demsetz 1982). The former approach implies that the existence of a barrier can in itself have no particular policy significance. The latter is policy relevant, but the complexity of most situations, combined with the subjectivity of the costs and benefits associated with proposed changes, makes it impossible to construct measures of barriers that are noncontroversial.

The next two sections of this paper trace the development of these ideas in the literature on entry barriers and discuss their relevance to the case of geographical mobility. The dilemma implied by these two approaches to barriers is illustrated in the fourth section, by reference to the impact on mobility of housing policy, and in the fifth section, by other contractual provisions that have the effect of restricting mobility. The final section suggests that the dilemma is ultimately irresolvable and that the term "barrier" has similar qualities to other items of economic terminology such as "subsidy" and "rent seeking," over which there is similar disagreement for substantially similar reasons.

Barriers in Economics

Bain (1956) provides the traditional starting point for a discussion of barriers to entry in economics. Starting from the observation that

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Professor of Economic Organization at the University of Buckingham.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

in competitive conditions entry will occur until price is equal to the average cost of production, Bain concludes that the persistence of prices above this level indicates the existence of entry barriers. Translated to the analysis of geographical mobility, Bain's conclusion would imply that barriers could be detected by observing, if it were possible, the rent accruing to labor.

Thus, suppose that Southern laborers were offered terms of employment that were much more favorable than those required to compensate Northern laborers for the marginal effort expended. Migration (i.e., entry) would be expected to compete away some of the rent accruing to Southern labor as "net advantages" were equalized between North and South. The persistence of differential returns to labor would then indicate a barrier. Because the relevant returns are not merely pecuniary but also include other highly subjective components of Adam Smith's "net advantages," we cannot infer the existence of barriers simply by observing wage differentials. But where physical restrictions on movement exist, queues of potential emigrants would be a measure of the size of mobility barriers.

Stigler (1968, p. 67) defines a barrier to entry as "a cost of producing . . . which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry." If instead of the new entrant we substitute the migrant, and instead of incumbent firms we think in terms of existing residents of an area, Stigler's definition of a barrier might be interpreted as including all the costs of physical movement such as transport, insurance, and search; the costs of buying and selling residential property; and so forth.

In the industrial organization literature, both Bain's and Stigler's approaches can lead to different conclusions about the sources of entry barriers. Economies of scale, for example, may constitute a barrier for Bain but never for Stigler, providing new entrants have access to the same technology as incumbents. In the field of geographical mobility, the two approaches are mutually compatible with Bain concentrating on the *effects* and Stigler on the *sources* of barriers. If migrants faced zero costs of change, no mobility barriers in the sense of Bain could persist.

Consider now the case of a local jurisdiction that decrees that everyone living in it should inhabit a house larger than a certain size and designed to a given set of specifications. Would such regulations constitute a mobility barrier? Given that the rules were to apply to established residents as well as to potential migrants, there would appear to be no barrier as defined by Stigler. Neither, in a world in which people could move costlessly, could the net advantages of living in the jurisdiction be greater than living elsewhere; thus, no

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

barrier would exist as defined by Bain. Clearly such regulations *would* affect the geographical distribution of the population, however. Certain (probably poorer) people who might otherwise have moved to the area will not be able to do so. Whether they can be said to face a barrier in the form of the local ordinances will depend on opinions concerning the desirability of the regulations. This criticism is essentially the one that Demsetz (1982) levels at the use of the term "barriers" in economics. By concentrating on the distinction between insiders and outsiders, Bain and Stigler "divert attention from other types of barriers" and "hide the value judgements implicit in the barriers notion" (Demsetz 1982, p. 48).

Although Demsetz does not provide a clear definition of a barrier, his approach clearly focuses on property rights. All systems of property rights imply barriers to entry: "The problem of defining ownership is precisely that of creating properly scaled legal barriers to entry" (Demsetz 1982, p. 19). Because there is a cost of transacting in property rights, social outcomes will be affected by how they are defined and allocated. A patent system, for example, may imply a barrier to the production of various goods, but the absence of a patent system would be a barrier to the introduction and discovery of new ones. Communal fishing rights imply a low barrier to entry into fishing but a high barrier to fish conservation. The licensing of taxis may be an entry barrier even when the licenses are tradable, but licensing might also be regarded as a barrier to unsafe, low-quality service.

From the point of view of public policy, the issue is whether existing property rights allocations are producing results that might be improved upon by some alternative system of barriers. Because the specification of a better alternative framework of property rights requires both clear predictions about the consequences of this postulated alternative state of affairs and a means of ranking one social state relative to another (i.e., a social welfare function), agreement about barriers is extremely difficult to achieve. Demsetz (1982, pp. 56–57) observes that the process of devising a framework of barriers is "rich in intuition and faith, and poor in discernible measurements. . . . Our utterances in this regard may be accorded the skepticism appropriate to fairly unadorned opinion."

In a series of papers, Lott (1986, 1987a, 1987b) has investigated the nature of barriers to entry and discussed their effects on the operation of political markets. Lott concentrates entirely on the question of the *transferability* of property rights. Barriers, he argues, exist when property rights are not tradable. For example, if barbers were licensed and if trade in the restricted number of licenses were not

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

permitted, a barber would continue to ply his trade, providing the return were sufficient to compensate him for the costs and effort involved. The existence of more efficient aspiring young barbers anxious to take his place would not affect the existing barber's retirement decision. With tradability, however, new and more efficient barbers can bribe the existing ones to sell their licenses at prices that leave both parties better off than they would have been in the absence of tradability. In other words, the absence of tradability creates a barrier to entry, mobility is inhibited, and an efficiency loss is added to the traditional loss associated with restriction of output equal to the difference in costs between existing and aspiring barbers. Similarly, Lott argues that because political reputations are person-specific and cannot be traded, new entrants into politics face a barrier, and incumbent politicians remain in office longer than would otherwise be the case.

Lott's approach is easily applicable to geographical mobility. Restrictions on the right to trade location-specific rights would constitute mobility barriers. Thus, suppose that I have a right to use a certain asset that cannot be moved from a given location. If I am not permitted to trade my rights in this asset, any decision to move will involve my sacrificing the present value of the flow of benefits derivable from the use of the asset. Clearly, I am less likely to move in these circumstances than I would be if I could sell my rights to this asset on the open market. The most celebrated example of this principle in action is the observed relative immobility of people with rights to use housing at below-market rents (e.g., Hughes and McCormick 1981).[1]

Transactions Costs and Mobility Barriers

Although there is no single definition of a "barrier" in the economics literature that commands universal assent, each approach mentioned in the previous section is related to the existence of transactions costs. Consider a world in which all property rights are exchangeable and the transactions costs associated with exchange are zero. Clearly, there are no mobility barriers in the sense of either Bain, Stigler, Demsetz, or Lott. In this frictionless world, all rents accruing to the inhabitants of an area would be instantly competed away by new entry. Insiders and outsiders would face the same cost

[1]Brittan (1977) was the first to argue that tenants in the public sector should be free to trade occupation rights as a solution to the mobility problem there. For the private sector, Ricketts (1986) puts the case for tradability of occupation rights as part of the process of deregulation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

conditions. If we invoke the so-called "Coase theorem," the final allocation of resources would be efficient, irrespective of the initial rights assignment (see Coase 1960) and no barriers as defined by Lott would exist by definition. It would seem reasonable, therefore, to focus on transactions costs as the ultimate source of barriers. A mobility barrier might then be defined as *any institutional arrangement that increases the cost of transacting in location-specific property rights.*

This definition highlights the unavoidable conceptual difficulties encountered when discussing barriers. The most important point is that any definition that refers to costs must implicitly be comparing what *is* with some alternative state of affairs. If a barrier is a "high" cost of transacting, we must face this question: "relative to what?" The analyst has various possible options.

First, an ideal and entirely fictional situation in which there are *no* costs of transacting might be used as a benchmark. This type of benchmark seems to be implicit in some of the work of Lott. It is not possible for one politician to sell his reputation for honesty, reliability, and service to someone else. Given that investments already made in establishing such a reputation are "sunk" costs, insiders and outsiders (those yet to make these investments) do not face the same cost conditions. Lott uses this observation to explain differences in campaign expenditures by challengers and incumbents. To describe the higher cost facing a challenger as a barrier in this situation, however, is to take as a point of comparison either a situation in which personal reputations are tradable at zero cost or are "creatable" at zero cost.

Second, instead of an imaginary zero transactions cost reference point, we might use some arbitrary, but realistically attainable, position. It might be argued, for example, that transactions costs will always exist and that our standard of comparison should be a situation in which public policy permits freedom of contract within a given framework of rights. The problem here is that there is no compelling logic behind the choice of any particular benchmark. The cost of transacting in location-specific rights depends on how they are assigned in the first place. Suppose, for example, that all the established residents in an area have to be compensated for loss of environmental amenities by any newcomer. Such a system will produce less mobility than a system in which the residents have to purchase the development rights from local landowners if they wish to conserve an area. In the former case, transactions costs impede mobility. In the latter case, they stand in the way of preserving the status quo.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Whether a mobility barrier exists will depend arbitrarily on which system we adopt as our point of reference.

Third, to provide a rationale for adopting a *particular* benchmark, the analyst may use a normative approach. A barrier becomes a cost of transacting in location-specific rights higher than would occur if property rights were defined and allocated in the best feasible way. This approach seems to represent Demsetz's view of the problem and explains, given the practical impossibility of identifying the optimum, his somewhat despairing opinion that barriers are almost completely subjective. Where one person sees a barrier, another sees the open road. Whether restrictions on the right to trade in slum housing, for example, represent a barrier to mobility will depend on opinions about the desirability of such restrictions. Similarly, processes of development control will be seen as inefficient mobility barriers by some people, but as efficient methods of environmental protection by local conservationists.

Location-Specific Rights and Housing Policy

Public policy in the field of housing has historically been associated with the creation of nontradable location-specific rights. At the end of the second section, I noted that below-market rents in the public and private sectors could be regarded as barriers to mobility if occupation rights were nontradable. The argument of the third section implies, however, (following Demsetz) that certain unstated normative judgments are hidden in this claim. In the public sector, for example, tradability of occupation rights would deprive housing authorities of control over the characteristics of their tenants. Whether this control is considered an important issue depends on opinions about alternative means of ensuring the availability of housing to relatively poor people, and upon opinions about the success or otherwise of the existing system in achieving socially desirable objectives other than maximum mobility of labor.

That restrictions on exchangeability will not always represent mobility barriers in the sense of Demsetz can be seen by considering cases in which they arise spontaneously in market settings. Consider the case of a housing cooperative or housing association. My rights as a member of a cooperative may not be exchangeable. Is this a mobility barrier? Clearly, if certain rights are costly or impossible to trade, I will be less inclined to move than otherwise would be the case. But in institutional settings, such restrictions on trade may be essential for achieving other purposes. Any private club that produces collective benefits, and not merely access, to a purely private

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

good is likely to restrict exchangeability in order to control the characteristics of new members. Willingness to pay the highest price may not be the appropriate criterion for choosing replacements. There may even be instances where the least-appropriate people are prepared to pay the highest price for membership, and nonexchangeability can be seen as a protection against adverse selection. In some regulated trades, such as medicine, where quality cannot be ascertained ex ante, this adverse selection problem has been used as a justification for restricting the exchangeability of licenses (e.g., Leland 1979). Thus, insofar as trade in the housing market encounters logically equivalent problems, restrictions on the tradability of property rights can be seen as both predictable and efficient.

If observed restrictions on tradability may be efficient, it is also true that the absence of restrictions is compatible with a barrier in the sense of Demsetz. In the owner-occupied sector, for example, it is possible to argue that mobility barriers are absent because there is a free market in this type of housing. The regional housing mobility index computed by Minford et al. (1987) is based on this assertion. Minford defines the mobility index to be 100 × (free-market rent ÷ actual rent) in the rented sectors and then sets the index equal to 100 in the owner-occupied sector. The argument is that under rent restrictions the market price of housing (i.e., the price that would exist if rent restrictions were removed) exceeds the actual price and, for the reasons discussed earlier, creates a barrier. The higher the index, the greater the barrier. The problem is that the free market position is not unambiguous. Suppose that development rights were held by all owners of property and no restrictions existed on residential development. In such a situation, it would be expected that the free market rent would be lower than that predicted under a system of rigorous planning controls. Thus, with the free-market rent calculated on this basis, the *lower* the mobility index, the *greater* the implied existing mobility barrier deriving from planning restrictions. There would be no presumption that mobility barriers were zero in the owner-occupied sector.

Other Sources of Barriers

In a world where information is not perfect or available on the same terms to everyone and, hence, where transactions costs are significant, institutions will often cope by *imposing* mobility barriers as defined by Lott. Incentives within a firm, for example, may involve employees accepting a rising profile of remuneration over time. This explanation of earnings profiles is discussed by Lazear (1981). A

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

rising profile implies that termination of employment will impose a cost on the employee. It is equivalent to a bond-posting device by which a person expresses confidence in his or her ability to do a good job by providing a hostage that will be forfeited in the event of poor performance. The whole purpose of the device is to induce loyalty and ensure the *durability* of a relationship that would otherwise be under threat from shirking and recontracting. It is a mobility barrier in the sense of Lott because it depends on existing workers not being able to sell to newcomers the value of their accumulated credibility or firm-specific reputation. If they wish to capitalize on past investments in good behavior, they simply have to stay where they are. Another similar example is the company pension, which is not portable in the sense that a change of employer greatly reduces the value of the pension. This device is clearly designed to discourage mobility and reward loyalty.

Some institutions, therefore, may require mobility barriers if they are to survive. Ban the company pension, insist that workers are paid entirely on observed performance at each point in time and not on the basis of seniority, and the result might logically be to render impossible certain hierarchical arrangements necessary for the production of particular goods. Economists are familiar with the basic proposition of a second-best welfare theory that technical conditions for efficiency, which are applicable to each and every market in an ideal world, do not continue to be applicable everywhere else when they are, for some reason, unavoidably contravened in one particular market. In a similar way, it appears that reducing mobility barriers is equivalent to a first-best policy and may not be appropriate where the costs of transacting can never be altogether avoided. In a world of transactions costs, in other words, artificially contrived mobility barriers may represent a second-best response.

Barriers and Economic Terminology

The difficulties that are associated with the concept of barriers in economics are encountered elsewhere and are not completely confined to this specific area. Indeed, there seems to be a class of terminology that is well established, yet is susceptible to logically similar problems.

First, the definition and measurement of a "subsidy," for example, requires a benchmark that is either normative (i.e., the subsidy implies a departure from some desirable alternative situation) or is arbitrary and, therefore, equally difficult to justify. In Ricketts (1985) I have given an example of the subsidy to owner-occupied housing.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The wide agreement among economists that there is such a subsidy depends on the consensus (to which I subscribe) that it would be desirable in the interests of economic efficiency if the rate of tax on owner-occupied housing were higher relative to the rate of tax on other types of housing or capital assets.

Second, another example of disputed terminology concerns "rent seeking." Usually this concept is defined in a normative way as "self-interested behavior which results in the use of resources in activities which are socially wasteful." As with the normative approach to barriers, this definition of rent seeking requires that we can identify alternative available situations that are not wasteful. Indeed, the barriers notion and the concept of rent seeking would appear to be closely connected since they both relate, when defined normatively, to situations in which property rights are inappropriately defined and assigned. Rents require barriers if they are not to be destroyed by new entry, and rent seeking may, therefore, be seen as investment in the creation of undesirable barriers or in the circumvention or sabotage of desirable ones. The lack of agreement characteristic of the debate about rent seeking (Ricketts 1987) can, therefore, be seen as rooted in precisely the same conceptual problems as the debate about barriers. As DiLorenzo (1988, p. 330) expresses the point, "one person's waste is another's cherished activity"—just as one person's barrier is another's justifiable protection. Virtually every kind of economic activity including research, development, and advertising (Cowling and Mueller 1978); efforts to pass examinations (Tullock 1980); hustling for legacies (Buchanan 1983); amateur rules in sport (Goff, Shughart, and Tollison 1988); political pressure (Pasour 1985); technological development in the fishing industry (Dnes 1985); and so forth have been described as rent seeking. In each case, the relevant analyst has argued that the system of barriers implicit in the established structure of property rights is inappropriate.

The term "barrier" figures prominently in discussions of many different areas of public policy. It is not the intention of this paper to argue that use of the term should be avoided or that it is inherently meaningless. However, it is very important to recognize the ambiguities that surround it and, in particular, to be aware of the overtly normative way in which it is very often used. In the context of the debate about rent seeking, DiLorenzo (1988, p. 319) argues not that economists can have nothing to say about the issue but that "making value judgements more explicit is likely to advance economic understanding." In the same way, arguments about barriers and the role of public policy in erecting or removing them will only lead to greater

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

understanding if protagonists are clear about the terminology they are using.

References

Bain, Joseph S. *Barriers to New Competition*. Cambridge: Harvard University Press, 1956.

Brittan, Samuel. *The Economic Consequences of Democracy*. London: Temple Smith, 1977.

Buchanan, James M. "Rent Seeking, Noncompensated Transfers, and Laws of Succession." *Journal of Law and Economics* 26, no. 1 (1983): 71–85.

Coase, Ronald H. "The Problem of Social Cost." *Journal of Law and Economics* 3, no. 1 (1960); 1–44.

Cowling, Keith, and Mueller, Dennis. "The Social Costs of Monopoly Power." *Economic Journal* 88, no. 4 (1978): 727–48.

Demsetz, Harold. "Barriers to Entry." *American Economic Review* 72, no. 1 (1982): 47–57.

DiLorenzo, Thomas J. "Property Rights, Information Costs, and the Economics of Rent Seeking." *Journal of Institutional and Theoretical Economics* 144, no. 2 (1988): 318–32.

Dnes, Anthony W. "Rent-Seeking Behaviour and Open-Access Fishing." *Scottish Journal of Political Economy* 32, no. 2 (1985): 159–70.

Goff, Brian L.; Shughart, William F.; and Tollison, Robert D. "Disqualification by Decree: Amateur Rules as Barriers to Entry." *Journal of Institutional and Theoretical Economics* 144, no. 3 (1988): 515–23.

Hughes, Gordon, and McCormick, Barry. "Do Council Housing Policies Reduce Migration between Regions?" *Economic Journal* 91, no. 4 (1981): 919–37.

Lazear, Edward P. "Agency, Earnings Profiles, Productivity, and Hours Restrictions." *American Economic Review* 71, no. 4 (1981): 606–20.

Leland, Hayne E. "Quacks, Lemons, and Licensing: A Theory of Minimum Quality Standards." *Journal of Political Economy* 87, no. 6 (1979): 1328–46.

Lott, John R. "Brand Names and Barriers to Entry in Political Markets." *Public Choice* 51, no. 1 (1986): 87–92.

Lott, John R. "The Effect of Non-Transferable Property Rights on the Efficiency of Political Markets." *Journal of Public Economics* 32, no. 2 (1987a): 231–46.

Lott, John R. "Licensing and Non-Transferable Rents." *American Economic Review* 77, no. 3 (1987b): 453–55.

Minford, Patrick; Ashton, Paul; and Peel, Michael. *The Housing Morass*. Hobart Paperback no. 25. London: Institute of Economic Affairs, 1987.

Pasour, E. C. "The Public Interest and Rent Seeking." *Public Choice* 47, no. 3 (1985): 525–29.

Ricketts, Martin. "The Subsidy as a Purely Normative Concept." *Journal of Public Policy* 5, no. 3 (1985): 401–11.

Ricketts, Martin. *Lets Into Leases*. Policy Study no. 81. London: Centre for Policy Studies, 1986.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Ricketts, Martin. "Rent Seeking, Entrepreneurship, Subjectivism, and Property Rights." *Journal of Institutional and Theoretical Economics* 143, no. 3 (1987): 457–66.

Stigler, George J. *The Organization of Industry*. Homewood, Ill: Irwin, 1968.

Tullock, Gordon. "Rent Seeking as a Negative Sum Game." In *Toward a Theory of the Rent-Seeking Society*, pp. 16–36. Edited by James M. Buchanan. College Station, Tex.: Texas A and M University Press, 1980.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Mutual Fund Banking: A Market Approach

Tyler Cowen and Randall Kroszner

Introduction

We examine mutual fund banking as an alternative form of financial intermediation. Individuals would hold checkable deposits at financial intermediaries structured as mutual funds. Although the nominal value of depositor holdings would not be fixed, risk could be hedged through the choice of portfolios. The regulations embodied in the Glass-Steagall Act, federal deposit insurance, and Federal Reserve operations would not be necessary to provide a sound and efficient banking and financial system. Our treatment of mutual fund banking is intended as a likely and viable scenario for the evolution of banking and financial institutions under laissez faire.

After a brief discussion of problems with current institutions, we outline the operation of a mutual fund banking system and the forces that would encourage its evolution and development. Comparisons and contrasts with a number of related ideas will help clarify the mutual fund banking proposal.

The Current Depository Institution Crisis

When examining policy alternatives, stability, and security of the monetary and financial sectors is a consideration of prime concern. Current regulatory policy utilizes deposit insurance and a lender of last resort as a response to the instabilities of a debt-based, fractional

---

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

Tyler Cowen is Associate Professor of Economics at George Mason University and Randall Kroszner is Assistant Professor of Economics at the University of Chicago, Graduate School of Business. The authors wish to thank Robert Barro, Richard Cooper, Kevin Dowd, David Glasner, Stephan Kalb, David Meiselman, Giovanna Mossetti, Joseph Salerno, and Stergios Skaperdas for useful comments and discussions. An earlier version of the paper was presented at the Cato Institute's Seventh Annual Monetary Conference, which was funded by the George Edward Durell Foundation.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

reserve banking system; such instabilities include bank runs and asset-liability maturity mismatches (see McCulloch 1981 and 1987).

Although deposit insurance has eliminated certain sources of financial instability, the cost has been high. The current deposit insurance system subsidizes bank risk and creates a moral hazard problem, ultimately at taxpayer expense. In the long run, deposit insurance is likely to make the banking system less secure, because depositors lose the incentive to monitor and foster stability (see Kane 1985).[1]

The regulatory dynamic has responded to the moral hazard problem associated with deposit insurance by regulating bank assets and bank risk taking. Such regulations, however, are not a long-term viable solution to the problem at hand. In the early 1980s, for instance, regulatory change was considered pressing because thrifts were rapidly losing money as a result of their inability to use asset diversification to insulate themselves from interest-rate risk. Maintaining previous regulations on thrift industry assets would also have created serious, although perhaps less-immediate, financial problems for the industry and the deposit insurance fund. Although the situation for commercial banks is not at a crisis level, as with the thrifts, regulators still face the Scylla of asset restrictions and capital requirements, which reduce bank profitability and viability against competitors, and the Charybdis of continued deregulation, which could engender moral hazard problems.

Current banking institutions and regulations are thus subject to long-run erosion. Mutual fund banking has the potential to alleviate many of the problems of traditional deposit banking—the combination of deregulation and mutual fund structure examined in this paper would radically differ from the current structure of depository institutions. Although depositors would not be required to use mutual fund banks, mutual fund banking is likely to arise in the absence of government intervention.[2]

[1]Deposit insurance has already created disastrous consequences in the savings and loan industry. Savings and loan institutions were allowed to diversity their asset holdings into riskier assets than mortgage loans, the traditional S&L staple. S&Ls subsequently invested in risky real estate ventures and loans, playing a "heads, I win; tails, the deposit insurance fund loses" game. As a result, government funding requirements may be as high as $300 billion.

[2]Recent advocates of variants of mutual fund banking include Black (1970), Fama (1980), Glasner (1987, 1989), Greenfield and Yeager (1983), and Kareken (1986); the related views of Simons (1948), Friedman (1960), and Litan (1987) will be discussed below.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Mutual Fund Banking

The alternative banking structure we explore entails changes on both the liability and asset side of bank balance sheets. Depository institutions would resemble mutual funds rather than current banks. On the liability side, the deposits of mutual fund banks would represent a claim on the asset portfolio held by the intermediary. Consequently, these bank liabilities may fluctuate in nominal value, depending on the worth of the underlying assets. "Non-par" banking would be possible, in which the number of deposit units or "shares" required to clear a check for a stated nominal sum would not be fixed at a one-to-one ratio. On the asset side, a wide variety of marketable claims, including equity, bonds, commercial paper, government obligations, options, futures, and commodities, would be available as investment vehicles for the fund.[3]

The ownership and management structure of mutual fund banks deserves special attention. The mutual fund intermediaries we discuss are different from the current-day mutual savings banks, credit unions, and mutual savings and loans, which we will refer to as mutual banks.[4] In both mutual banks and mutual funds banks, the depositors constitute the bank's "shareholders." One becomes an "owner" of the bank by making a deposit and ceases to be an owner only by withdrawing the deposit. In the mutual fund banking proposal we consider, depositors are residual claimants and hold a direct claim to the assets of the intermediary, thus earning a return tied to the performance of the underlying portfolio. Under the historical structure of mutual banks, in contrast, depositors receive a predetermined interest payment and are not true residual claimants.

[3]Although mutual fund bank accounts would not have a fixed par value, mutual fund banks might find it convenient to maintain par value check clearing as an accounting convention for the course of ordinary business. Fluctuations in the nominal value of the account would still occur, but the value of a share would be priced at one, and capital gains would simply increase the number of shares. A check written out for a certain number of dollars would be used to transfer the same number of shares; a similar procedure is often used for today's money market mutual funds. If an account suffers sufficiently large capital losses, however, par value clearing might no longer be sustainable. The purchasing power of each share would then float with the price of the underlying asset basket, and the account would be marked to market each time a check was cleared.

[4]Rasmusen (1988) and Woodward (1988) analyze these institutional forms. Rasmusen shows that mutual banks held a significant percentage of time deposits (sometimes more than 80 percent) before the onset of deposit insurance and also had significantly lower failure rates than shareholder-owned banks. Rolnick and Weber (1988) discuss a form of mutual fund banking that prevailed during the American free banking era. See also Goodhart (1987), Selgin and White (1987), and Wicker (1987) on issues related to mutual fund banking.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Mutual fund banks are likely to evolve along the lines of current money market mutual funds. Depositors would invest in open-end mutual funds in which depositors/shareholders would be full residual claimants to the performance of the asset portfolio.[5] The mutual fund would be managed by a separate entity that probably would be a division of a "financial supermarket," a publicly traded stock or private company such as J. C. Penney, Dreyfus, Merrill Lynch, or Citibank. As with today's mutual funds, managers could be compensated with a percentage of the value of the fund's assets, although more complex performance-based reward schemes are also possible.[6]

Depositors could select their portfolios from a large set of possible alternatives. For the most risk averse, funds holding purely short-term federal government or government-guaranteed issues would provide the highest degree of security from capital value fluctuation, as both interest-rate and bankruptcy risk would be negligible. Other investment-grade instruments of varying maturities, combined with the use of interest-rate futures, also would have a high degree of safety.

For securities or commodities with well-developed and liquid options and futures markets, various hedging strategies are available that can be used to eliminate much of the risk of capital value fluctuations. The stock index futures market, for instance, permits hedging of broad equity portfolios. A great diversity of risk-return configurations—for example, insuring downside risk without eliminating upside potential—are available through existing markets for securities and their derivatives.[7]

The development of mutual fund banking would undoubtably augment the liquidity of these rapidly developing markets and encourage additional innovation. Individuals need not learn the intricacies

[5]A closed-end mutual fund format could be used also, although such funds often trade at a premium or discount to their net underlying asset value (see Thompson 1978). The unresolved questions surrounding closed-end fund pricing make them less likely candidates than open-end funds.

[6]There is thus no danger that "depositors' funds" would be used to subsidize nonbank activities; the mutual fund principle requires the bank to serve as a pure intermediary. Account holders and not banks would possess the legal claims to the assets behind their accounts; the only "funds" remaining for the use of the intermediary would the commissions charged on accounts.

[7]In the absence of deposit insurance, the use of bank debt (rather than equity) as a form of wealth insurance for depositors appears counterintuitive, as insurance theory implies that individuals wish to insure their portfolio against large swings in value, not small swings. Fixed nominal claims imply that thinly capitalized banks (themselves vulnerable to large losses through runs) are protecting depositors against small nominal losses. Insurance policies, in contrast, usually protect against large losses but not small losses (i.e., they contain a deductible). See Calomiris and Cone (1984).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of futures and options markets to hedge; mutual fund banks would offer a variety of claims to fully, partially, and unhedged funds, so the depositor need only choose his desired risk-return structure (see, e.g., Fama 1980). Hedging, of course, cannot decrease the aggregate risk faced by the economy as a whole, but it can shift this risk away from depository institutions and the payments system.

The superior returns available on certain financial instruments would be attractive for those who do not demand full guarantees and are willing to accept a higher variance of returns. Equity portfolios, for instance, outperform other investments with respect to long-run rates of return.[8] In addition, even strongly risk-averse depositors may prefer asset holdings with varying nominal returns and will not necessarily seek refuge in fixed nominal claims that do not protect the real value of the holder's assets. The loss of personal savings through inflation is a common theme in economic theory; as recently as the last 1970s, Americans were earning negative real rates of return on interest-bearing banking accounts.

Bank Stability

In contrast to traditional banks, depository institutions organized upon the mutual fund principle cannot fail if the value of their assets declines. Since the liabilities of the mutual fund bank are precisely claims to the underlying assets, changes in value are represented immediately in a change in the price of the deposit shares. The run-inducing incentive to withdraw funds at par before the bank renders its liabilities illiquid by closing vanishes with the possibility of non-par clearing. In effect, there would be a continuous (or, say, daily) "marking to market" of the assets and liabilities. Such a system obviates the need for much of the regulation long associated with a debt-based, fractional reserve system, as the equity-nature of the liabilities eliminates the sources of instability associated with traditional banking institutions.

Although money market mutual funds operate without government insurance, their safety record has so far been impeccable. With assets in excess of $300 billion, they have gained the confidence of a large set of depositors (*Federal Reserve Bulletin*, January 1989, p. A45). Money market funds have been able to handle very large and rapid inflows and outflows without disruption. In 1983, for instance, money market mutual fund shareholders withdrew roughly one-quarter of

[8]Smith (1988) reports that in the 1926–87 period stocks delivered an average annual nominal return, including dividends, of 9.9 percent. In comparison, bonds averaged 4.9. Mehra and Prescott (1985) present a detailed, technical study of superior stock returns—the so-called equity premium puzzle.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

their total balances without precipitating any runs or liquidity problems (McCulloch 1987, p. 8). A drain of such proportion from bank deposits for cash is unlikely to have been so smooth. Withdrawals of a smaller percentage were enough to trigger the banking crises of the early 1930s (see Friedman and Schwartz 1963, pp. 712–13). While widespread demands for redemption may depress the value of depositor assets, the solvency of mutual fund banks cannot be threatened by such redemptions.

Even panic selling in equity markets will not disrupt the stability of mutual fund banks. An equity market panic would involve a sale of asset claims, a fall of asset prices on the secondary market, and a capital loss for depositors; however, no further first-order effects on financial intermediation need occur. In contrast, a run on a debt-based bank may require the bank to call in loans (thus interrupting production), or liquidate loans. Traditional bank assets are typically illiquid, and the rush to liquidate will impose further costs on the bank.[9]

Historical experience illustrates the importance of preventing the bankruptcy of financial intermediaries. Benjamin Bernanke (1983), for instance, provides strong evidence that the bankruptcy of financial intermediaries was a significant factor contributing to the depth and persistence of the Great Depression. Intermediary bankruptcy disrupts the flow of capital from savers to borrowers, which chokes off investment, disrupts long-term plans, and leads to business bankruptcies. The breakdown of the intermediation and credit allocation system, according to Bernanke's evidence, appears to have made the Great Depression "Great."

Commercial Lending

With the transformation of banking institutions to a mutual fund basis, commercial loans now originated through banks instead would be financed largely through commercial credit facilities and securities issues. Even currently, where banks receive the benefits of membership in the Federal Reserve System and have access to federally insured depositors' funds, other financial intermediaries have competed successfully with banks in the commercial loan market. In the early 1950s, for instance, banks' share of lending to (nonfinancial) corporations was nearly 90 percent; by 1985, this figure had fallen to just slightly above 60 percent. As of 1985, banks held less than 40

[9]Dowd (1989) and Cowen and Kroszner (1989) discuss the use of the option clause, which permits banks to suspend redemption temporarily and pay interest, to mitigate the causes and consequences of runs in debt-based, fractional reserve systems.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

percent of the collective assets of financial institutions (Litan 1987, pp. 45–46).

Bank competitors include diversified financial "supermarkets" such as Merrill Lynch, and finance subsidiaries of major companies, such as Ford and General Motors, which lend to home purchasers as well as automobile purchasers. General Electric's finance subsidiary, GE Credit, is now one of the 10 largest commercial lenders in the United States (Litan 1987, p. 188). Such commercial lenders fund their loans by issuing equity and debt securities, rather than using deposits, thus illustrating that the efficient allocation of funds does not require the linkage of deposit taking and commercial lending.[10]

Commercial credit lending does not simply remove the runs problem to another environment (as argued by Diamond and Dybvig 1986). In contrast to a bank with (debt-type) demand deposits, the commercial lender has control over the maturity structure of its liabilities. Unlike demand deposits, commercial credit liabilities, such as commercial paper and long-term debt, cannot be presented for redemption at their holder's will. Equity shares in the lending firm, of course, cannot be presented for redemption either. These liabilities of the commercial credit firm can still supply liquidity (even with later maturities) if they are traded in secondary markets. Furthermore, a commercial credit lender can fail without inducing a run on the banking system.

The rapidly expanding commercial paper market has been another important bank competitor in commercial lending, as major corporations increasingly have bypassed banks and gone directly to the securities markets for debt financing. From 1960 to 1985, outstanding commercial paper increased fortyfold to over $200 billion, whereas bank loans increased only elevenfold over the same period to $494 billion. For large manufacturing corporations (those with assets over $1 billion), commercial paper accounts for 52 percent of their short-term borrowing; bank loans account for only 26 percent (Litan 1987, p. 42). As the costs of communications and monitoring technology continue to fall and financial markets continue to grow more sophisticated, more and more companies, not only the largest, will find the

[10]The process of "securitization," repackaging illiquid loans (e.g., auto, credit card, mortgage) into marketable securities traded in liquid markets, provides another rapidly growing alternative to depositors' funds for financing illiquid loans. Rosenthal and Ocampos (1988, p. 221) estimate that there is already $100 billion in outstanding securitized credit without any government guarantees, with another $600 billion in government-guaranteed securitized mortgages. Greenbaum and Thakor (1987, p. 397) argue that securitization (without government guarantees) would be more prevalent if there were no deposit insurance or subsidies offered through federal credit agencies.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

commercial paper market a feasible and desirable means of debt finance.[11]

Eliminating the Glass-Steagall and other regulatory barriers to entry into the underwriting of corporate securities would expand the securities financing possibilities for firms of all sizes, especially small firms that are on the margins of the corporate securities market. In addition, smaller firms that may not have access to the commercial paper and equity markets would still have access to loans and funding from the commercial credit lenders and capital pools. Commercial credit and investment banking firms could open local branches in search of smaller-scale lending opportunities, just as many large banks have done recently (Litan 1987, p. 181).

Comparisons and Contrasts with Related Proposals

Henry Simons (1948) and Milton Friedman (1960) have suggested 100 percent reserve banking to mitigate the instabilities and inefficiencies associated with a fractional reserve banking system.[12] Such proposals are motivated by many of the concerns we have voiced in this essay. One hundred percent reserves against demand deposits would provide an alternative to deposit insurance and address the problem of bank runs. Simons and Friedman also would permit free entry into deposit creation services, as long as assets of approved types were held as the 100 percent reserve.[13]

In contrast to mutual fund banking, demand deposits in the Friedman-Simons plan would be fixed nominal value claims. While Simons briefly considered the mutual form as a viable alternative for savings accounts, the Friedman-Simons plan did not explore the possibility of structuring demand deposits as mutual fund shares. In addition, the Friedman-Simons plan allows only a narrowly defined set of assets to be used as reserves. The reserves would be issued by the Treasury or central bank and, in Friedman's version, would earn interest. Even with interest payments on these funds, however, banks might wish to circumvent the government-mandated portfolio and hold a higher-yielding asset portfolio. The constraints on the

[11]Note that commercial paper issuers often obtain a backup "letter of credit" from a bank or other source that promises to pay if the issuing company cannot.

[12]Many writers have advocated various forms of 100 percent reserve banking. See Kareken (1986) for a recent discussion of this approach and Diamond and Dybvig (1986) for a critique.

[13]Simons and, somewhat more ambivalently, Friedman also made government monopoly of the currency supply part of the proposal. In contrast, we envisage mutual fund banking as a move toward currency competition and away from government control of high-powered money.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

earnings of depository institutions under the Friedman-Simons proposal would raise the costs of using bank services; as a result, depository institutions might attempt to offer bank-line services while avoiding the legal label of "bank."

Robert Litan (1987) has put forth a proposal for "narrow banking" that shares some features with the mutual fund approach outlined here. Litan's "narrow banks" would take deposits and receive federal deposit insurance but would not engage in personal or commercial loans; those functions would be carried out by separate subsidiaries of a financial holding company. The separately incorporated lenders would face few or no restrictions on their assets but would neither receive federal deposit insurance nor have access to the Fed's check-clearing and payments system. The assets held by narrow banks would be limited to highly liquid and "safe" securities, as determined by the regulatory authorities. Capital requirements might range from fractional reserves to collateralization in excess of 100 percent, and insurance premiums also would be set by the regulatory agencies. Litan (1987, pp. 169–170) envisions a guaranteed nominal value for narrow bank deposits.

Despite some parallels, mutual fund banking institutions would not require the complex regulatory oversight that Litan's narrow banking proposal would require. By treating depositor balances as mutual fund shares subject to non-par clearing, the possibility of bank failure is eliminated, and there is thus no need for deposit insurance and the regulatory structure it entails. Under Litan's proposal, in contrast, narrow banks would not be strong marketplace competitors but would survive only insofar as regulatory protection prevented other intermediaries from entering the payments system. Litan (1987, p. 185) acknowledges that there is an "end-run" problem (which also plagues the Friedman-Simons proposal): Non-insured intermediaries, which can circumvent the asset restrictions, may issue deposit-like liabilities yielding a better return than that offered by the narrow banks. If the public is willing to accept and to use these as demand deposits, then the narrow banks' liabilities are dominated and will not be held, undermining the proposal.

Optimal Banking Structures: The Roles of Equity and Debt

Over the last decade, a large and growing theoretical literature has focused on the problem of optimal financial institutions and their regulation (see Williamson 1987 for a survey). The results of this literature are motivated by some form of asymmetric information and

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the costs of monitoring. The bank, for example, may not know the actual pay-off of a borrower's project without costly auditing. In other cases, "preference shocks" to depositors' demand for withdrawals may be known only individually and may lead to welfare-reducing bank runs (see Jacklin and Bhattacharya 1988).

Models such as Townsend (1979) and Diamond (1984) find that diversification arguments and information asymmetries concerning project payoffs imply that financial intermediaries with fixed-commitment, debt-like liabilities are optimal. Focusing more on the depositors' conveniences, transactions costs and risk-sharing arguments for the optimality of a debt form for demand deposits also are put forth by Goodhart (1987, pp. 86–87) and Woodward (1988). Diamond and Dybvig (1983), drawing from Bryant (1980), argue intermediaries perform the key role of creating liquidity by transforming illiquid assets (e.g., loans) into liquid liabilities (e.g., standard debt-liability demand deposits) (see also Freeman 1988). In a model with some restrictions on trading opportunities (e.g., no credit cards), Jacklin (1987) shows that debt contracts provide better risk sharing than equity, although this advantage disappears if the restrictions are removed.

In the influential Diamond and Dybvig (1983) model with standard debt-type demand deposits, banks may be subject to runs, even when the banks' assets are riskless. Government deposit insurance, as well as alternatives to the standard demand deposit contract, can eliminate the runs problem. In a modification of the Diamond-Dybvig framework, Postlewaite and Vives (1987) find the even stronger result that bank runs are the unique equilibrium with fixed-commitment deposit contracts.

More recent developments in the theoretical literature demonstrate that the optimality of the debt contracts is not robust to reasonable changes in the theoretical model. By generalizing the Townsend approach to consider a richer set of monitoring technologies, most importantly random auditing, Mookherjee and Png (1987) conclude that state-contingent claims such as equity contracts are optimal and debt never is. In a model with asymmetric information about bank asset quality, Jacklin and Bhattacharya (1988) find that debt or equity may be optimal, depending upon the nature of the information available about the assets.

Unlike some forms of 100 percent reserve banking (e.g., the Friedman-Simons plan), mutual fund banking realizes the gains from trade resulting from the transformation of illiquid assets into liquid liabilities. Mutual fund banks still perform the valuable service of transforming mutual fund shares into liquid claims for transactions use.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The existence of a check and clearinghouse system enables these assets to be used for making purchases for current consumption, thus increasing the liquidity of the community's assets as a whole. Furthermore, the liabilities of commercial credit firms used to fund loan assets could be liquid as well, if traded in secondary markets. Maturity matching of assets and liabilities does not rule out the provision of liquidity services. For this reason, the institutional structure of mutual fund banking would have no problem performing the traditional "liquidity transformation" functions of banks.

In general, while the specific environments of these formal banking models may be tenuous bases for policy conclusions, a mutual fund banking system is consistent with recent developments in the theory of optimal financial intermediation.

The Evolution of Mutual Fund Banking

The current structure of banking institutions stems in large part from legal restrictions on intermediation and banking services (see, e.g., Hall 1982 and Cowen and Kroszner 1988). If different financial intermediaries were allowed to compete on equal terms, mutual fund banking might evolve through the competitive pressures of the risk-return preferences of depositors and the desire of intermediaries for secure and profitable enterprise.

The potential for mutual fund banks is illustrated by the structure of nonbank financial intermediaries. When individuals invest money outside the network of depository institutions, non-par-value contracts such as stocks, mutual funds, and money market funds are common.[14] Certificates of deposit (CDs) are the primary fixed value investment claim that successfully competes with these non-par-value instruments. CDs, however, can offer only competitive rates of return by restricting access to depositors' funds, an infeasible option for deposit banking.[15]

Money market mutual funds have continued to flourish even after restrictions on bank interest rates (e.g., Regulation Q) were removed (Lown 1987). Furthermore, even with the advent of money market deposit accounts, where money market rates of return are offered

[14]Mutual insurance companies are also very common. Companies with a mutual ownership structure possess almost two-thirds of the total assets of the life insurance industry and account for one-half of the total amount of life and health insurance in the United States (Huebner and Black 1982, p. 567).

[15]Even with many CDs, the depositor's rate of return depends on the state of the world, such as the performance of the stock market. See Rankin (1988, p. 7). Of course, even CDs function as non-par-value instruments when they are "liquified" through sale at non-par-value in a secondary market.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

through banks with deposit insurance protection, uninsured money market mutual funds have remained competitive. Money market mutual funds have been successful despite regulatory obstacles that put them at a disadvantage relative to "depository institutions" vis-à-vis services of the Federal Reserve.[16]

Increasing the scope of competition from mutual fund depository institutions would require a number of fundamental policy changes. First, Congress would have to repeal the provisions of the bank regulatory statutes that prohibit banks from freely purchasing, underwriting, and dealing in equity. Second, the Federal Reserve System could allow mutual funds and money market funds to enter its check-clearing and wire transfer systems, or privatize these systems and allow them to be supplied competitively. More generally, current fears of predatory pricing or regulation in retaliation for building a network to compete with the Federal Reserve's clearing system provide powerful disincentives against private innovation and investment in this area. Third, the tax-induced biases against equity and for debt would also require modification. The current structure of capital gains taxes also discourages mutual fund banking by requiring that every transfer of equity instruments be recorded for the purposes of eventual taxation.

Level competition among financial intermediaries also would require the dismantling of deposit insurance. The FDIC, for instance, could gradually lower the amount of bank deposits that are covered by insurance. The current $100,000 limit could be lowered $10,000 each year, thus phasing out deposit insurance over a 10-year period. Banks with highly liquid asset portfolios (like today's money market mutual funds) would gradually attract depositors from banks that held riskier portfolios of commercial loans. As these latter banks lost deposits, their loan portfolios could be purchased by commercial credit lenders, or perhaps the entire bank would be taken over by nonbank intermediaries.

Conclusion

Depository institutions should be evaluated in terms of their ability to withstand volatile economic conditions and crises as well as by their success in serving the credit and transactions demands of the economy; mutual fund banks appear to offer considerable promise on these issues. Mutual fund banking thus merits serious consider-

[16]A list of Fed services accessible to only "depository institutions" can be found in *Leveling the Playing Field* (1984, p. 14). See also Calomiris and Cone (1984).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

ation as a policy alternative to current banking institutions and as a model for the evolution of a deregulated financial sector.

References

Bernanke, Benjamin. "Nonmonetary Effects of the Financial Crisis in the Propagation of the Great Depression." *American Economic Review* 73 (June 1983): 257–76.

Black, Fischer. "Banking and Interest Rates in a World Without Money." *Journal of Bank Research* (1970): 9–20.

Bryant, John. "A Model of Reserves, Bank Runs, and Deposit Insurance." *Journal of Banking and Finance* 4 (December 1980): 335–44.

Calomiris, Charles, and Cone, Kenneth. "A Note on Competitive Payment Systems." Unpublished manuscript, Northwestern University, 1984.

Cowen, Tyler, and Kroszner, Randall. "The Evolution of an Unregulated Payments System." Unpublished manuscript, Harvard University, 1988.

Cowen, Tyler, and Kroszner, Randall. "Scottish Banking before 1844: A Model for Laissez-Faire?" *Journal of Money, Credit, and Banking* 21 (May 1989): 221–31.

Diamond, Douglas. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51 (1984): 393–414.

Diamond, Douglas, and Dybvig, Philip. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (1983): 401–19.

Diamond, Douglas, and Dybvig, Philip. "Banking Theory, Deposit Insurance, and Bank Regulation." *Journal of Business* 59 (January 1986): 55–68.

Dowd, Kevin. "Monetary Freedom and Monetary Stability." Paper presented at the Cato Institute's Seventh Annual Monetary Conference, Washington, D.C., 1989.

Fama, Eugene. "Banking in the Theory of Finance." *Journal of Monetary Economics* 6 (January 1980): 39–57.

Freeman, Scott. "Banking as the Provision of Liquidity." *Journal of Business* 61 (January 1988). 45–64.

Friedman, Milton. *A Program for Monetary Stability*. New York: Fordham University Press, 1960.

Friedman, Milton, and Schwartz, Anna J. *A Monetary History of the United States, 1867–1960*. Princeton, N.J.: Princeton University Press, 1963.

Glasner, David. "Competitive Banking Without Deposit Insurance." *Cato Policy Report* 9 (March/April 1987): 1, 10–12.

Glasner, David. *Free Banking and Monetary Reform*. New York: Cambridge University Press, 1989.

Goodhart, Charles. "Why Do Banks Need a Central Bank?" *Oxford Economic Papers* 39 (1987): 75–89.

Greenbaum, Stuart I., and Thakor, Anjan V. "Bank Funding Modes: Securitization Versus Deposits." *Journal of Banking and Finance* 11 (1987): 379–401.

Greenfield, Robert, and Yeager, Leland B. "A Laissez-Faire Approach to Monetary Stability." *Journal of Money, Credit, and Banking* 15 (August 1983): 302–15.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hall, Robert. *"Monetary Trends in the United States and United Kingdom:* A Review from the Perspective of New Developments in Monetary Economics." *Journal of Economic Literature* 20 (December 1982): 1552–56.

Huebner, S. S., and Black, Kenneth. *Life Insurance.* Englewood Cliffs, N.J.: Prentice Hall, 1982.

Jacklin, Charles. "Bank Deposits, Trading Restrictions, and Risk Sharing." In *Contractual Arrangements for Intertemporal Trade.* Edited by Edward Prescott and Neil Wallace. Minneapolis: University of Minnesota Press, 1987.

Jacklin, Charles, and Bhattacharya, Sudipto. "Distinguishing Panics and Information-Based Runs: Welfare and Policy Implications." *Journal of Political Economy* 96 (June 1988): 568–92.

Kane, Edward. *The Gathering Crisis in Federal Deposit Insurance.* Cambridge, Mass.: MIT Press, 1985.

Kareken, John. "Federal Bank Regulatory Policy: A Description and Some Observations." *Journal of Business* 59 (January 1986): 3–48.

Leveling the Playing Field: A Review of the DIDMCA of 1980 and the Garn–St. Germain Act of 1982. Chicago: Federal Reserve Bank of Chicago, 1984.

Litan, Robert. *What Should Banks Do?* Washington, D.C.: Brookings Institution, 1987.

Lown, Cara S. "Money Market Deposit Accounts versus Money Market Funds." *Economic Review,* Federal Reserve Bank of Dallas (November 1987): 29–38.

McCulloch, J. Huston. "Misintermediation and Macroeconomics Fluctuations." *Journal of Monetary Economics* 8 (1981): 103–15.

McCulloch, J. Huston. "The Ohio S&L Crisis in Retrospect: Implications for the Current Federal Deposit Insurance Crisis." Working Paper, Department of Economics, Ohio State University, 1987.

Mehra, Rajnish, and Prescott, Edward. "The Equity Premium: A Puzzle." *Journal of Monetary Economics 15* (1985): 145–62.

Mookherjee, Dilip, and Png, Ivan. "Optimal Auditing, Insurance,and Redistribution." Unpublished manuscript, Stanford University, 1987.

Postlewaite, Andrew, and Vives, Xavier. "Banks Runs as an Equilibrium Phenomenon." *Journal of Political Economy* 95 (June 1987): 485–91.

Rankin, Deborah. "The Puzzling New World of Bank C.D.'s." *New York Times,* 16 October 1988, p. F7.

Rasmusen, Eric. "Mutual Banks and Stock Banks." *Journal of Law and Economics* 31 (October 1988): 395–422.

Rolnick, Arthur J., and Weber, Warren E. "Explaining the Demand for Free Bank Notes." *Journal of Monetary Economics* 21 (1988): 47–71.

Rosenthal, James, and Ocampo, Juan. *Securitization of Credit: Inside the New Technology of Finance.* New York: John Wiley & Sons, 1988.

Selgin, George, and White, Lawrence H. "The Evolution of a Free Banking System." *Economic Inquiry* 25 (1987): 439–58.

Simons, Henry. *Economic Policy for a Free Society.* Chicago: University of Chicago Press, 1948.

Smith, Randall. "Why Buy?" *Wall Street Journal,* 19 October 1988, p. 1.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Thompson, Rex. "The Information Content of Discounts and Premiums on Closed-End Fund Shares." *Journal of Financial Economics* 6 (1978): 151–86.

Townsend, Robert. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* 21 (October 1979): 265–93.

Wicker, Elmus. "Is Non-Par-Value Deposit Banking a Solution to the Bank Failure Problem?" Unpublished manuscript, Indiana University, 1987.

Williamson, Stephen D. "Recent Developments in Modeling Financial Intermediation." *Quarterly Review*, Federal Reserve Bank of Minneapolis (Summer 1987): 19–29.

Woodward, Susan. "A Transaction Cost Analysis of Banking Activity." *Cato Journal* 7 (Winter 1988): 683–99.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The Precarious Nature of Sovereign Lending: Implications for the Brady Plan

Melanie S. Tammen

Introduction

The so-called Brady Plan—the new policy of officially subsidized reduction of the private debts of less developed countries (LDCs), as announced by U.S. Treasury Secretary Nicholas Brady on March 10, 1989—is merely a new twist in the U.S. government's eight-year-old policy of preventing LDC debtors and their foreign commercial creditors from working out soured LDC debts on their own terms.[1] This U.S. policy of avoiding conventional-style default at all costs originated with the Treasury Department's and the Federal Reserve's crisis management of Mexico's 1982 moratorium on principal payments to its foreign creditor banks. Given that U.S. money center banks' loan exposure to Latin America in 1982 averaged 160 percent of their capital base, and that the U.S. government (that is, taxpayers) were already leashed through federal deposit insurance to the bank's plight, there was little alternative to the initial interventionist policy.

In 1990, however, unlike eight years ago, European, U.S., and Japanese banks are increasingly well positioned to formally realize the imbedded losses on their LDC loans. Indeed, through innovative, privately initiated, and mutually beneficial debt-reduction techniques, the creditor banks and LDC debtors retired nearly $50 billion in external LDC debt from 1984 through early 1989—before the Brady Plan. The Brady Plan, in fact, chilled much of this private

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Director of the Cato Institute's Global Economic Liberty Project.

[1]While the foreign debt of all LDCs totals $1.3 trillion, the foreign debt of the "rescheduling" LDCs—those enmeshed in the debt crisis of the 1980s—totals about $600 billion.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

debt-reduction activity, particularly LDC debtors' debt-equity swap programs.

In the guise of "jump-starting market-based debt reduction," the Brady Plan is only a new development in the post-1982 U.S. policy of papering over LDCs' de facto (and de jure) defaults on foreign bank loans. This policy continues to allow private banks in the United States and other industrial countries to collect more interest—and suffer fewer losses—on their bad LDC loans by slowly transferring much of the bad debt exposure to the multilateral lending agencies, and thus to the industrial nations' taxpayers. Proponents of the Brady Plan argue that it will be a cost-effective endeavor if the commercial banks absorb large losses in the officially subsidized, debt-reduction agreements. As the argument goes, debtor governments will gain so much breathing room that they will dutifully service all obligations on the remaining debt. In short, once debt is restructured, in many cases with World Bank and IMF interest payment guarantees in place, the LDC debt problem will vanish. Like the "war to end all wars," the Brady Plan is to be the rescheduling to end all reschedulings.

Such a scenario defies not only the experience of hundreds of years of sovereign debt defaults, but also the post-1982 record, whereby major debtors have each rescheduled and stretched out their foreign commercial bank debts between three and nine times. For centuries governments have borrowed private money, squandered it on corrupt and unproductive state schemes, defaulted, and then negotiated easier terms with their creditors, only to default again. The fundamental aim of the Brady Plan, like the previous debt policy of U.S. Treasury Secretary James A. Baker, is to prevent the *private* bargaining process whereby indebted sovereign borrowers (the honest and dishonest alike) and their creditors have long settled delinquent debts at a fraction of their nominal value.

Debtors and creditors would exploit debt-reduction techniques of the private sector far more aggressively if not for LDCs' annual receipt of large World Bank and IMF loans. Throughout the mid- to late 1980s, the indebted LDC governments effectively "round-tripped" these funds to private creditor banks as debt service. Also since the late 1980s, the World Bank and IMF appear to be increasingly round-tripping much of their disbursements to borrowers back to themselves, as debt service requirements on old World Bank and IMF obligations grow. In a protracted attempt to paper over de facto LDC defaults with new credits, U.S. policy under Secretary Brady, as under Secretary Baker before him, is permanently chaining LDC debtors to financial and economic micromanagement by the World

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Bank and IMF. In the course of this process, the policy is financing the build-up of the next LDC debt bomb, which will involve World Bank and IMF debts.

Previous episodes of sovereign debt default suggest that the debts which LDC governments acquired during the 1970s and 1980s also will eventually be settled at a fraction of their face value. In the present debt crisis, the official debts to the international lending agencies, which are steadily replacing the bad private debts, will go largely unpaid. Yet, previous episodes of sovereign default, particularly the 1930s episode, indicate that debtors and creditors can, indeed, negotiate significant levels of debt reduction when left to work things out for themselves and to exploit the development of a secondary market in debt claims.

Centuries of Sovereign Default: A Brief Review

The precarious nature of sovereign lending is preserved in a history of default that dates back to medieval times (Lewis 1986/87).[2] The English monarchy developed such a record of default that creditors learned to require tangible collateral. Edward III, for example, had to pawn the crown jewels to obtain Florentine bank credit for his part in the Hundred Years War. The famous German banking house of the Fuggers helped bankroll the early success of the Hapsburgs in expanding their dominion over Germany, the Netherlands, Belgium, Italy, the Iberian Peninsula, and beyond. The Fuggers' problems in collecting on those loans drove them from the sovereign lending business.

The most spectacular defaulter in the early history of international debt was Philip II of Spain. He turned to the Genoese bankers after the Fuggers declined to finance his various enterprises, which included the Armada, the reconquest of Belgium, and campaigns against the Ottomans. While Spain's New World minerals brought unprecedented wealth, Philip's finances were in perpetual disarray. (This chaos was perhaps the first example of a government's natural resource windfall abetting an unsustainable fiscal outlay.) Various types of bonds were invented to fund Philip's imperial projects and these bonds were traded actively, with large discounts emerging when Philip's credit rating was low. In March 1577, for example, the paper traded for 55 percent of face value. Philip defaulted on his Genoese debt in 1557, 1560, and 1576. He then turned to Portuguese financiers, and defaulted on debt to them in 1590, 1606, and 1627.

[2]Except as otherwise noted, this section relies heavily on Lewis's (1986/87) study, to which Tammen contributed.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

The structure of international lending remained largely unchanged from the 17th through the early 20th centuries. Foreign lending took the form of either short-term trade finance or special purpose loans at fixed interest rates. The latter were used to build much of South America's and Russia's infrastructure. Sovereign debt defaults remained common, particularly throughout the 19th century. Creditor country governments generally maintained a hands-off approach, with certain notable exceptions, such as the British occupation of Egypt in 1880 and the French occupation of Mexico in 1863. But these exceptions were justified on political grounds, not on the basis of collecting bad debts. Even the U.S. gun-boat diplomacy in the Caribbean from the 1890s through the 1920s was based more on protecting the Panama Canal than on protecting private investors.

The interwar years were marked by sovereign defaults throughout Central and Eastern Europe, China, and Latin America. This period was also marked by some highly resourceful strategies on the part of borrowers to benefit from their default. When Germany had difficulty meeting obligations stemming from its World War I reparations, the United States organized various teams of experts to negotiate easier terms. The 1924 Dawes Plan stretched reparations out over a 50-year period—with payments tied to a prosperity index—and included a special loan to help halt hyperinflation. Still, as soon as Germany had to make significant net repayments to the United States, it defaulted. The subsequent 1929 Young Plan further stretched out reparations and provided another special loan, but this plan was eventually defaulted on as well. Brandeis University's historian Stephen Schuker (1988) examined the period's balance of payments statistics and found that the net capital flow ran to Germany during both the inflation and stabilization phases of the Weimar Republic. Not only did the Reich entirely avoid paying net reparations to its wartime opponents; it actually extracted the equivalent of reparations from the Allied powers, principally the United States. The gross capital inflow amounted to 5.3 percent of German national income during the entire period of 1919–31.The net capital inflow, after subtracting all reparations transferred and after making generous allowance for the disguised return of German funds, still came to a minimum of 2.1 percent of national income over the same period (Schuker 1988, pp. 10–11).

These reparations to Germany allowed the maintenance of living standards in the Weimar Republic at a level appreciably higher than was justified by domestic production. The Weimar experience provides a striking parallel to Latin American governments' postponement of economic adjustment throughout the 1980s, while those

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

governments annually contracted for large levels of new loans from the World Bank and IMF. As Schuker (1988, pp. 11, 46) writes:

> Savings and investment remained notably low compared with either the prewar pattern or the long-term trend. The inflow of funds accommodated increased wages and salaries, even in sectors with lagging productivity gains, and despite the more precipitous decline in the length of the work week in Germany than elsewhere. These funds also found reflection in mounting government welfare expenditures before as well as after the onset of the Depression, in an uneconomic shift to white-collar employment in labor-force composition, and (although precise figures remain a matter for conjecture) in the accretion of German assets abroad that would later help finance Nazi rearmament. . . . As the Depression deepened, Germany possessed a rationalized industrial plant, a social welfare system unrivaled anywhere except in Great Britain, and a conglomeration of municipal amenities that commanded the wonder of the world. The foreigners who had financed much of it held paper claims—just as they had in 1919–23.

Finally, the round of Latin American bond defaults of the 1930s is most relevant to the debt crisis of the 1980s. The international commercial upheaval of 1920–21 brought a sharp drop in export revenues, which resulted in large public deficits and balance-of-payments problems for Latin American governments. To refinance many loans, as well as to finance ambitious new public works and urban modernization projects, most Latin American governments embarked on a foreign loan splurge. Between 1922 and 1928, U.S. banks sold almost $2 billion in Latin American bonds to North American investors (Marichal 1989, pp. 182–84).

By the late 1920s, however, a decline in bond quality showed that investors realized Latin American borrowings were being used to meet earlier bond obligations. After 1929, most Latin American nations were unable to float new issues. By 1930, only a third of all Latin issues was being serviced in full. In short, the disappearance of new money preceded most defaults. The situation was similar to Weimar Germany, which was willing to repay past loans only so long as the net transfer of capital was favorable.

In relating the Weimar and Latin American defaults to the present round of de facto defaults, a liquidity squeeze as precipitator of default is only the first of several parallels. In recent years, numerous proposals have surfaced for a global solution to LDC debt, and they strikingly resemble plans floated in the 1930s (Eichengreen and Portes 1989, p. 80). Six or more "debt facility" proposals, such as one introduced in 1988 by American Express Chairman James D. Robinson III, would have governments of creditor countries finance

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

a new facility—likely linked to the IMF or World Bank—that would buy up LDC debt at a discount and negotiate with debtors for disposition of the debt.

Similarly, the Kindersley-Norman Plan of 1931 (named after a director and the governor of the Bank of England) would have had governments of the main creditor countries capitalize a new international entity that would finance itself by selling bonds to private investors. Another proposal floated in the 1930s would have given the Bank for International Settlements (established in 1930 as part of the Young Plan) the resources necessary to intermediate and resolve the bond crisis. These plans required a serious commitment by the governments of major creditor countries, namely the United States and United Kingdom. Such commitments did not materialize; large investors, too, were uninterested. J. P. Morgan argued that investors would not fund any international facility unless control over it would rest with them (Eichengreen and Portes 1989, p. 81).

While the various global plans never got off the ground in the 1930s, privately initiated transactions retired a significant level of debt, chiefly through Latin governments' repurchases of their bonds in the secondary market at prices well below par. In 1939, the Foreign Bondholders Protective Council[3] estimated that a dozen countries in default had repurchased 15–50 percent of their bonds since 1930. Erika Jorgensen and Jeffrey Sachs (1989) estimate that Bolivia repurchased 5 percent of its defaulted debt at an average price of 16 cents on the dollar and that Chile retired 18 percent at 59 cents; Colombia, 22 percent at 22 cents; and Peru, 31 percent at 21 cents on the dollar.

In public, bondholder committees opposed the repurchases and insisted that available foreign exchange should be used to service the bonds on contractual terms. But the committees were more receptive to the practice in private than in public, especially when the repurchase was accompanied by resumption of at least partial debt service payments, and also when the creditors were unlikely to receive a better offer (Eichengreen and Portes 1989, p. 82). In 1937, Britain's Corporation of Foreign Bondholders (CFBH), according to its records, quietly acquiesced in Chile's use of $4 million for repurchases. In 1940, as part of a settlement recommended by the CFBH, Brazilian authorities devoted "at least $400,000 in each of four years" to purchases in the British market. And in 1941 negotia-

[3]This council was founded by U.S. bondholders in 1934, with State Department encouragement, and was patterned after Britain's Corporation of Foreign Bondholders, which was founded in 1868.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tions, the CFBH grudgingly agreed to Colombia's request that it be allowed to purchase all of its British debt in the market.

Even more revealing were bondholders' responses to efforts of the U.S. Securities and Exchange Commission to curb the repatriation of bonds by defaulters by restraining banks and brokers from dealing on behalf of governments in default. CFBH records note that during SEC hearings in 1937, bondholder representatives maintained that such restraints on repurchases would be met with "strong and . . . effective criticism on the grounds that, by limiting the market in such bonds, it would act detrimentally to the bondholders" (Eichengreen and Portes 1989, p. 82).

Nearly 50 years ago, Henry Wallich (1943, p. 332) admitted the ethical problem that "arises when repurchases are made after the bonds have depreciated owing to suspension of service, for in that case the repurchasing debtor is profiting from his own default." Yet, Wallich (1943) also noted a great advantage to repurchases of Latin American government bonds in light of the high export earnings of such countries during World War II: "If part of the reserves that are currently being acquired [by debtor countries] are not used for repurchases now, the chances are that after the war they will be utilized for imports and not for the service of foreign debts."

Wallich, of course, proved to be correct. When settlements were finally negotiated, British and American creditors did recover their principal, on average.[4] But the realized internal rate of return for several issues reached substantial negative levels (for example, −7.4 percent for Brazil, −9.8 percent for Bolivia, and −14.76 percent for Hungary), indicating that not just interest but substantial principal was written off (Eichengreen and Portes 1989, p. 79).

The benefits that Brazil and Mexico derived from protracted settlements suggest that their present governments should take a long view of the resolution possibilities for their external debt difficulties. Brazil suspended interest payments on most of its foreign debt in 1931. In 1932, a plan was announced to issue 20- and 40-year bonds to capitalize interest arrears, with interest payments to resume in 1934. In 1934, however, a set of easier terms was negotiated. In

[4]Calculations by Eichengreen and Portes (1989, p. 79) suggest that, for creditors with diversified foreign bond portfolios who were willing to hold out for final settlement, the average nominal internal rate of return (weighted by issue value) was roughly 4 percent on dollar bonds and about 5 percent for sterling issues. Although dollar bondholders settled for approximately half of contractual interest and sterling bondholders settled for only slightly more, dollar bondholders did only slightly worse than if they had held domestic Treasury bonds, and sterling bondholders did slightly better. Ex ante risk premiums were nearly sufficient to compensate American bondholders for foreign lending risks, and the premiums more than sufficed for British bondholders.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

1937, debt service payments were suspended again. In 1940, Brazil announced a temporary settlement. Finally in 1943, Brazil and its creditors negotiated a permanent readjustment, with creditors suffering losses as noted above (Eichengreen and Portes 1989, p. 78).

As for Mexico, only when it received massive U.S. wartime assistance in 1942 was there a lasting agreement on its foreign bonds. These bonds were late 19th-century government borrowings and railway bonds on which Mexico first defaulted in 1914. A succession of debt agreements followed throughout the 1920s and 1930s. One such agreement in 1930 consolidated all Mexican government bonds into a single, new, 45-year bond with a submarket interest rate and all interest arrears wiped out. In 1931, debt service was paid in blocked pesos; and in 1932, it was suspended. Then in 1942, despite the fact that the projected flow of U.S. aid alone would have facilitated full service of Mexico's obligations, claims of $510 million were paid off for less than $45 million—roughly nine cents on the dollar (Lewis 1986/87, p. 50).

The 1980s Antidefault Cartel

In the German and Latin American defaults of the 1930s, permanently assembled international lenders of last resort, such as the World Bank and IMF, did not exist.[5] As related above, sovereign debtors and their foreign bondholders negotiated debt reschedulings and settlements directly, usually through bondholder committees. The Weimar Reichsbank's President Hjalmar Schacht and his Latin American counterparts recognized the Roosevelt administration's general indifference to bondholders' concerns and learned to exploit divisions among their creditors (Schuker 1988, p. 140). While these debtors profited from their own delinquency, to the extent that taxpayers in creditor nations remained largely out of the loop, this solution was relatively equitable.

In the 1980s round of sovereign defaults, creditors were chiefly commercial banks from Europe, the United States, and Japan. By definition, banks are far more integrally tied into national and international monetary systems than are bondholders. Yet, despite the fact that banking systems of Europe, the United States, and Japan are already backed by national lenders of last resort, the key feature

[5]By a strict definition of lender of last resort, the World Bank and IMF do not qualify given that they do not have the authority to create money (that is, to provide unlimited liquidity upon demand). However, because both institutions lend to numerous countries that cannot attract such sums in the marketplace, they bear a significant resemblance to lenders of last resort and are, thereby, described here as such. For further discussion, see Barth and Keleher (1984).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

of the U.S. LDC debt policy since 1982 has been the formation of a veritable cartel of lenders of last resort to prevent conventional default on sovereign debts at all costs. In response to the Mexican payment moratorium in 1982, Federal Reserve Chairman Paul Volcker pressed the IMF to leverage its lending power by coercing Mexico's creditor banks into involuntary lending as the price for preserving their claims.[6] From this action, the cartel quickly grew to encompass the Federal Reserve and industrial nation (OECD) central banks, the IMF, the multilateral development banks (chiefly, the World Bank), the Paris Club of official creditors, the U.S. Treasury, the Group of 7 finance ministers, and the U.S. Federal Deposit Insurance Corporation. Anna Schwartz (1988, p. 15) describes Washington's socialization of private loans to LDCs succinctly:

> The strategy devised by the U.S. treats not only the debtor countries but also the creditor banks as wards of the U.S. regulators. . . . The regulators abetted the accumulation of the debt by U.S. banks, praising them for effectively recycling surplus current account funds of OPEC countries, and abstaining from cautionary injunctions as the portfolio of foreign debt grew.[7] The regulators permitted the banks to evade the provision that limits loans to a borrower to no more than 15 percent of a bank's capital. Loans to various state-operated firms and institutions in a foreign country—all guaranteed by the government—were regarded as individual loans.
>
> When the debt problems erupted, the banks were not urged to reduce dividends and build loan loss reserves. Instead public policy was based on the fiction that the Latin American debts on the books of the banks were assets that had full face value, despite their discounted value in secondary markets.

In the early 1980s, U.S. bankers and regulators considered the capitalization of debts in arrears (while preserving the book value of existing obligations) as altogether different from default. Indeed, Jeffrey Sachs (1982, p. 226) argued: "In no sense is private debt rescheduling merely a polite name for default." While, historically, bond reschedulings had seldom resulted in debtors eventually hon-

[6]In response to Mexico's August 1982 suspension of principal payments on its external debt, the U.S. Treasury Department and the Federal Reserve Board, in conjunction with the Basle-based Bank for International Settlements, moved quickly to provide $1.85 billion as an emergency bridge loan to Mexico. In addition, the United States provided $1 billion in food aid and another $1 billion in prepayment for Mexican oil. The Federal Reserve was chiefly responsible for securing for Mexico a $3.6 billion IMF loan and a $5 billion loan from Mexico's foreign creditor banks. This was the first instance of what would afterward become the standard practice of linking a new IMF loan with a rescheduling (often with new money) agreement by a commercial bank.

[7]Here, Schwartz references statements by Chairman Burns (1977) and Chairman Volcker (1980).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

oring their obligations in full after having been "tide over" with easier terms or new funds, the U.S. debt strategy aimed to overcome this record. Still, it must be granted that the severely over-exposed position of U.S. banks to Latin America at the time of the 1982 Mexican moratorium—combined with government backing of U.S. banks through federal deposit insurance—meant that the initial official U.S. response could hardly have been otherwise.[8]

In October 1985, U.S. Treasury Secretary Baker updated the tide-them-over strategy, moving from an approach on the basis of macroeconomic austerity to one in which the IMF, World Bank, and private creditor banks would provide debtors with substantial net flows in exchange for market-oriented reforms. The premise was that debtors would reform forthwith and grow their way out of debt. The Baker Plan called for LDC debtors, over a three-year period, to receive $9 billion in loans from multilateral development banks (MDBs) and another $20 billion in loans from private bank creditors.

The Baker targets were, in fact, met. From 1985 through 1988, the "Baker 15" middle-income LDCs that were heavily indebted received $20.2 billion in disbursements from their private creditor banks and $22.8 billion in disbursements from multilateral creditors (including $15.9 billion from the World Bank) (Wertman 1990, p. 8). Over this three-year period, both commercial banks and official creditors managed to maintain positive net flows (disbursements on new loans minus amortization on old loans) to the targeted group of LDCs. Official creditors averaged $6 billion annually; commercial banks averaged $3.3 billion annually. But in strict terms of cash flows or net transfers (disbursements on new loans minus debt service on old loans) during the period, official creditors averaged only $700 million annually, and commercial banks averaged $-17 billion annually (World Bank 1989, p. 37).

The wide divergence in private versus official new loan provision has effected a slow transfer of relative LDC debt exposure from the commercial banks to Western taxpayers by using international institutions. Since 1982, LDC debt held by official creditors (the IMF, World Bank, and industrial nation governments) has grown by 135 percent, while LDC debt held by private banks has increased by only 46 percent (World Bank 1989, p. 2).

[8]The Federal Reserve reports that in 1983 exposure of U.S. banks, as a percentage of capital, to all LDCs was 172 percent (271 percent for U.S. money center banks); exposure to Latin America, in particular, was 106 percent (all U.S. banks) and 160 percent (U.S. money center banks).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Despite commercial banks having met the Baker Plan's target for new lending during 1985–88, the negative net transfers to the banks indicate that banks were lending only those sums required to ensure continued repayment on previous loans. Baker's early 1988 announcement of support for a $75 billion capital increase for the World Bank ($14 billion to come from the United States) was borne of the necessity to have the World Bank continue annual positive net transfers to the debtors to counteract for private banks' negative net transfers. Just as the World Bank's capital increase was an integral part of making the Baker Plan work, the November 1989 announcement of U.S. support for a funding increase for the IMF[9] was synonymous with making the new Brady Plan work. The Brady Plan contained a key role for the IMF; and the price for the IMF Managing Director Michel Camdessus's playing ball was the capital expansion he had quietly sought for several years.

During the mid-1980s, economic commentators first began referring to the Ponzi scheme (or pyramid scheme) that was represented by commercial banks extending to LDCs only those new loans that were necessary to get paid on old loans. Somewhat later, in 1988, LDC debt watchers began decrying a new development in the Ponzi scheme—the round-tripping of World Bank (and IMF) loans to Western commercial banks as debt service. Round-tripping was accomplished largely through the World Bank's relatively new policy-based lending, which in recent years has packaged as much as $500 million in a single loan for quick disbursement to the largest Latin debtors.[10] On June 13, 1989, for example, the World Bank approved three $500 million policy-based loans to tide Mexico over further.

Since 1989, a further development in the Ponzi scheme has come into view. As of April 30, 1989, IMF loans that were outstanding to borrowers and were more than six months in arrears totaled nearly

[9]In May 1990, the IMF's member governments, including the United States, formally initialed a 50 percent increase in IMF resources, or quotas, from $120 billion to $180 billion. The increase required a new contribution of $12 billion from the United States.

[10]The World Bank's policy-based lending (that is, loans in exchange for promises of economic reform rather than for infrastructure projects) has been stepped up since 1980 and particularly following the 1985 initiation of the Baker Plan. Mexico, Argentina, and Brazil have each received one or more $500 million policy-based loans in recent years, for such things as agriculture sector reform or trade sector reform. Such loans are ostensibly used to provide the hard currency needed to finance LDC imports, but in fact they are nothing more than general budgetary support. While all money is fungible, this nonproject-related assistance is highly fungible. Indeed, the amount of a policy-based loan (or package of them) is determined solely by the size of a debtor nation's financing gap in a given period—a clear signal that the funds are largely used for private debt service.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

$4 billion. Since 1986, IMF borrowers in this category increased from 8 to 11, and overdue obligations increased six-fold. Growing accounts in arrears have afflicted the World Bank, too, where $3.2 billion in loans are on nonaccrual status. These growing numbers of World Bank and IMF loans in arrears indicate that the World Bank- and IMF-financed Ponzi scheme is coming home to roost. The two lending goliaths now have to consider what annual level of new loans is necessary so that even they can get paid.

Table 1 illustrates this borrowing treadmill, vis-à-vis the World Bank, on which 17 highly indebted middle-income countries are now stuck. Despite new loan commitments to these countries being up from $6.1 billion in 1986 to $8 billion in 1989, net transfers from the World Bank have fallen from $319 million to $ − 1.9 billion over the period. The World Bank, linchpin of the Baker—and now the Brady—strategy, is caught in a trap of dumping ever-higher levels of new loans into the indebted LDCs while its negative net cash flow to them widens.

Finally, a deteriorating bottom line at the IMF is also signaled by its new forbearance framework, which loosely parallels the relaxed regulatory environment that proved so fateful to the U.S. savings and loan industry during the 1980s: an environment of taxpayer guarantees. In September 1989, IMF Managing Director Camdessus confirmed a formal policy change whereby the IMF will now extend loans to debtor countries that had not yet, at that time, reached agreements with their private creditors (Camdessus 1989, p. 8). Since May 1989, the IMF has approved more than $10 billion in loans (Philippines, Costa Rica, Mexico, Venezuela, and Argentina) under this framework, which forwards new loans to LDCs that are able to pry ever-less money out of their private creditors. While at first glance this transfer may not appear so critical—all the affected countries but Argentina did conclude agreements with their private creditors relatively soon after the IMF loans—it indicates that as LDCs' private bankers increasingly pull the plug, the IMF and the World Bank are opening the spigot. Under the Brady debt strategy, the extent to which the World Bank and IMF will be forced to round-trip funds to themselves will only accelerate.

In fact, the Brady Plan's role for World Bank and IMF guarantees on restructured, written-down commercial bank debt—already a part of the 1990 debt-reduction agreements of Mexico and Venezuela—will only increase the pressures on the World Bank to maintain positive annual net transfers to the affected LDCs. Particularly with the World Bank, the fear that even one activation of an LDC's interest guarantees might jeopardize the Bank's AAA credit rating will be

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 1

THE WORLD BANK AND THE 17[a] HIGHLY INDEBTED, MIDDLE-INCOME COUNTRIES: GROWING COMMITMENTS AND DECLINING NET TRANSFERS ILLUSTRATE THE PONZI SCHEME

(US$ MILLIONS; FISCAL YEARS)

	1986	1987	1988	1989	Total 1986–89
IBRD and IDA Commitments	6,070.5	6,719.4	6,482.8	8,021.0	27,293.7
Gross Disbursements	4,212.8	6,132.1	5,405.5	4,739.9	20,490.3
Repayments	1,907.5	2,709.4	3,492.2	3,546.3	11,655.4
Net Disbursements	2,305.3	3,422.7	1,913.1	1,193.6	8,834.9
Interest and Charges	1,986.4	2,646.6	3,180.7	3,119.2	10,932.9
Net Transfer	318.9	776.1	−1,267.4	−1,925.6	−2,098.0

[a]The 17 countries include Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cote d'Ivoire, Ecuador, Jamaica, Mexico, Morocco, Nigeria, Peru, Philippines, Uruguay, Venezuela, and Yugoslavia.

SOURCE: World Bank (1989a, p. 47).

enough to keep the World Bank annually committing substantial new loans to these borrowers.

Real-Politik among Private Bankers and Sovereign Debtors

Despite the fact that U.S. policy until late 1987 stubbornly retained the fiction that LDC debts were worth 100 cents on the dollar,[11] bankers began facing reality well before then. Starting in 1982, a secondary market in LDC debt slowly developed as banks undertook interbank swaps to restructure lopsided sovereign portfolios and to manage risk. Market volume received a boost when Brazil began allowing debt-equity swaps in 1984, and Chile and Mexico instituted the first formal debt-equity programs in 1985. Total secondary market transactions increased from $2 billion in 1984 to $50 billion in 1988. Within this market, annual debt conversions arranged by commercial banks plus LDC public and private debtors grew 10-fold, from $2.1 billion in 1985 to $22.4 billion in 1988 (see Tables 2 and 3).

The leading instrument for debt conversion has been the debt-equity swap, in which an investor purchases debt in the secondary market at a discount and exchanges it for local currency, bonds, or state-owned equity shares from the debtor government. In recent years, the potential benefits of debt-equity swaps for LDC economies have been largely left out of the popular discourse on LDC debt. Debt-equity swaps have been unfairly criticized as necessarily inflationary and as a subsidy to direct investments in LDCs, which would have taken place in any event. Besides whittling away at LDC debt, debt-equity swaps also (a) attract new foreign investment; (b) attract back-home "flight capital" held abroad by LDC citizens; (c) provide long-term financing for LDC companies when domestic credit markets are tight; and (d) frequently finance new export-oriented investments, which earn much-needed hard currency.

The Additionality Debate

LDC officials frequently complain that debt-equity swaps lack "additionality" (that is, they do not pull new foreign investment funds into debtor countries, but merely subsidize investments that would have been made in any event). Joel Bergsman and Wayne

[11]At the September 1987 annual meetings of the World Bank and IMF, Treasury Secretary Baker first announced the updating of his debt strategy with the menu approach, which supported such debt relief options as exit bonds and debt-equity swaps. This strategy appears to mark the first time the U.S. government recognized that the foreign debts of numerous LDCs were not worth face value.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 2

DEBT CONVERSIONS BY SELECTED COUNTRIES, 1984–89

(US$ MILLIONS)

Country	1984	1985	1986	1987	1988	1989[a]
Debt Conversion[b]						
Argentina	31	469		35	1,130	500
Brazil	731	537	176	1,800	9,175	4,000
Bolivia				1	349	20
Chile	11	313	987	1,983	2,905	2,000
Costa Rica			7	146	17	10
Ecuador				125	258	
Honduras				6	11	
Jamaica				2	100	
Mexico		769	1,023	3,804	6,670	6,000
Nigeria					95	200
Peru					15	
Philippines			43	287	806	300
Sudan					1	
Uruguay					97	
Venezuela					477	
Yugoslavia					50	
Zambia					3	
Other						750
Total	773	2,088	2,236	8,188	22,358	13,780
All Transactions[c]	2,000	4,000	7,000	12,000	50,000	40,000

[a]Projected.
[b]Debt-equity, debt-debt swaps (that is, transactions that convert external debt claims on the debtor country to another form), and debt repurchases. Debt-equity swaps are recorded at the book value of debt converted into equity.
[c]Debt swaps; all transactions including interbank transactions; estimates.
SOURCE: World Bank (1989b, p. 15).

Edisis (1988) of the World Bank's affiliate International Finance Corporation (IFC) interviewed the investing participants of 104 debt-equity transactions in an effort to understand the extent of additionality. Bergsman and Edisis (1988, p. 10) found that the swap mechanism made a difference in nearly half of the swaps by multinational corporations, as well as in all of the swaps arranged by creditor banks for their own accounts—a total of 61 percent of the swap transactions studied.

The IFC study also found that additionality increases as swap programs mature. Since foreign investments are one to two years in

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 3

CONVERSIONS OF DEVELOPING COUNTRY EXTERNAL DEBT, 1984–88

(US$ MILLIONS)

	1984	1985	1986	1987	1988
Conversions					
Debt-Equity Swaps	773	1,843	1,522	3,335	9,205
Exit Bonds	0	0	0	15	4,725
Buy-Backs	0	0	0	0	648
Informal Swaps	0	0	0	3,500	5,414
Other	0	245	714	1,337	2,366
Total Conversions	773	2,088	2,236	8,188	22,358

SOURCE: World Bank (1989a, p. 18).

gestation, most investments in the early stages of a swap program are ones that would have occurred without the subsidy. As more and more investors became aware of the benefits of a program, these benefits become decisive for a larger percentage of transactions. For example, investors from many parts of the world, including countries that have not had close business ties with Chile, have started to look at Chile for investment possibilities. Since 1987, largely through debt conversion, Chile has attracted investments of $200 million from Australia, $250 million from South Korea, $250 million from Taiwan, and $600 million from New Zealand (Brooke 1989).

Additionality is also higher for investments in export-oriented industries, which earn much-needed hard currency for the host country.[12] The IFC study identified several cases in which swaps caused businesses to create new export-oriented companies in Latin America rather than in Southeast Asia. In other cases, investments in pre-existing companies soon led to increases in capacity and to startups of entirely new lines of production. Table 4 gives further evidence that debt-equity swaps are an important catalyst for sorely needed foreign direct investment; it indicates that such investments in Mexico and Chile have been highest during the period when debt-equity swaps have been in use.

[12]Conversely, investments oriented toward domestic markets tend to be those that would have been made without the swap incentive.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

TABLE 4

FOREIGN DIRECT INVESTMENT (FDI) THROUGH SWAPS AND BEFORE SWAP PROGRAMS: CHILE AND MEXICO

(US$ MILLIONS; ANNUAL AVERAGES)

	FDI through Swaps (1985–87)	Normal FDI (1979–82)	Depressed FDI (1983–86)
Chile	400	139	65
Mexico	1,133	791	141

SOURCE: Bergsman and Edisis (1988, p. 3).

Debt-Equity Programs: Chile's Model, Others' Disappointments

Largely through its debt swap program, Chile has retired approximately $8 billion in foreign debt since 1985—equivalent to more than half of its medium- and long-term commerical bank debt at year-end 1985. Chile's debt service ratio (the annual debt payments as a percentage of export revenues) fell to 28 percent in 1988 from 73 percent in 1982 (*Reuters* 1989). In short, Chile has essentially conquered its foreign debt crisis.

Chile's 1985 debt conversion program has been well detailed in numerous sources;[13] only a brief reiteration of its model features is necessary. To prevent its swap program from fueling inflation, Chile's central bank honors the debt presented by investors with 15-year inflation-indexed bonds rather than pesos. The investor obtains cash by selling the bond in Chile's capital market. This arrangement, together with the central bank's monthly ceilings of $100 million in Chapter 19 conversions (relating to foreign investors), sterilizes the potential inflationary effect. Chile's 1988 inflation rate was a manageable 12 percent—tame by Latin American standards. To avoid granting the swap subsidy to all foreign investments, the rules for capital and profit repatriation are stricter under Chile's swap program than the rules covering foreign investments made with new money outside the mechanism. Chilean nationals, who participate through debt-for-local-currency swaps, accounted for nearly half of the swaps completed through mid-1989 ($2.4 billion). Foreign investors accounted for $2.6 billion (Durr 1989).

Critical to the Chilean program's success has been the maintenance of a very favorable investment climate, produced in large part by the government's decisive attack on the fiscal deficit and its

[13]See, for example, Hanke (1987) and Tammen (1989).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

sustained and aggressive privatization program. Chile's economic growth has averaged about 6 percent over the past three years. The Chilean debt-equity program's straightforward design and administration also have been critical to its success.

While nearly a dozen LDCs, aside from Chile, had some type of debt-equity swap program in place in 1988, most programs have included terms that deter potential investors. Many programs have been intermittently restricted or suspended. Some LDCs have weighed down their swap programs with new money requirements—stipulations that an investor must bring new funds into the country outside the swap program to finance a portion of the investment. For example, Argentina's October 1987 program includes a 30 percent new money requirement. In addition, in Argentina swaps cannot be used to purchase existing enterprises, including state-owned companies. This latter restriction is particularly regrettable given that each year Buenos Aires has to find 30 percent of the total financing costs of 117 state industries (Mead 1989).

In Brazil, the debt-equity program precludes the acquisition of a majority foreign interest in a Brazilian entity. In the Philippines, the central bank charges a 20 percent fee on swaps, effectively halving the discount that the investment receives. In 1988, Manila's central bank, fearing the swaps' inflationary impact, set a $180 million per year ceiling on debt swaps. It also announced that new investments would receive preference for approval over investments in existing facilities. The various restrictions served to nearly halt debt-equity activity in the Philippines. In Mexico, the government halted its 18-month-old debt-equity program for public-sector debt in November 1987 ($3 billion had been retired), citing its inflationary impact and concerns that the swaps were subsidizing investments that would have taken place without the incentive.[14] Still, throughout 1988 and early 1989, private Mexican firms acted on their own initiative to directly buy back their debt from foreign banks with cash at steep discounts. By year-end 1989, these so-called corporate restructurings

[14]International automobile manufacturers, including Chrysler Corp., Ford Motor Co., Nissan Motor Co., Ltd., and Volkswagen AG, accounted for much of the initial activity under the Mexican debt-equity swap program. Since the firms were planning expansion prior to the program, Mexican officials have often cited these swaps as evidence of a lack of additionality. The planned expansions were, in fact, in response to an early 1986 Mexican government decree that forced foreign automobile firms to step up exports. In addition, as mentioned earlier, the first swaps in any program will be linked to investments previously under consideration.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

had retired nearly all of Mexican private firms' $14 billion in foreign debt.[15]

Other Private Debt Reduction Innovations

In 1988, Chile and Bolivia recalled how the widespread debt buy-backs, which followed the 1930s bond defaults, allowed the debtor to capture fully 100 percent of the secondary market discount. Bolivia repurchased $240 million of its foreign debt—not serviced since 1984—at only 11 cents on the dollar with funds anonymously donated by foreign governments. Chile spent $168 million to buy back and retire $299 million of its foreign debt, paying an average 56 cents on the dollar. In September 1989, Chile announced plans to buy back another $550 million of its remaining $6.3 billion in outstanding commercial bank debt.[16] Today, the chief advantage for participating commercial banks remains what Wallich noted after the bond defaults of the 1930s. Straight debt buy-backs offer the prospect for debtor countries to divert some hard currency funds away from wasteful domestic spending or the financing of import consumption to at least partial debt repayment.

In early 1988, under a mechanism pioneered with Morgan Guaranty Trust Company, Mexico had hoped to use $2 billion in reserves to purchase U.S. Treasury bonds worth $10 billion at maturity in 20 years. These bonds would collateralize the principal of $10 billion worth of new government bonds that Mexico would issue and exchange with its foreign commercial creditors for $20 billion in current Mexican debt—effecting principal forgiveness of 50 percent. But banks were unhappy that interest on the new bonds would carry no collateral backing. Projected interest payments were estimated at about 85 percent of the total flow of funds to bond holders. Creditor banks bid average discounts of 30 percent. Mexico exchanged $2.6 billion in new bonds for $3.7 billion in old Mexican debt—achieving debt reduction of only $1.1 billion. Afterward, the Mexican government initiated efforts to arrange a new version of the debt-for-bonds swap, which would carry World Bank or creditor government guarantees on interest. The 1989 Bray Plan contains just such a provision, which has subsequently been incorporated into the 1990 debt-reduction agreements of Mexico and Venezuela.

[15]Conversation with Peter Truell of the *Wall Street Journal*, January 4, 1990.

[16]Accordingly, in November 1989, Chile purchased another $140 million in commercial bank debt at a discount of 41 percent. The hard currency funds for this buy-back, however, came from World Bank and IMF borrowings under the framework of the Brady Plan.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Heavier Doses of Reality: Creditor Banks Reserve for Losses

In 1983–84, commercial banks in Europe initiated a first round of discrete provisioning for expected losses on LDC loans, under direction from their banking authorities. These provisions were tax deductible in most European countries. In contrast, U.S. regulators encouraged banks to maintain the 100-cents-on-the-dollar fiction. In fact, in 1985 Reagan Administration officials proposed to eliminate the small tax incentive that encouraged U.S. banks and thrifts to provide for anticipated loan losses. The partial tax deductibility of loan loss reserves ended in 1986, and such reserves became deductible only at the time of charge-off—when losses are actually recorded. At the same time, however, loan loss reserves did count toward meeting U.S. banks' and thrifts' regulatory capital, despite their providing no protection to the insurer once bad debts are properly written off (therefore, such reserves did not constitute a capital cushion.) Throughout the 1980s, most other creditor country regulators, including those in Canada, Germany, the Netherlands, Switzerland, and the United Kingdom, helped to strengthen their banks' balance sheets by providing tax deductibility for loan loss reserves, while not allowing reserves to count as capital. By contrast, U.S. regulators have had it backward.

By early 1987, the secondary-market prices of LDC loans had so depressed the prices of U.S. bank stocks that the over-exposed banks no longer had much to lose by initiating orthodox loan loss provisioning. In May 1987, Citicorp put aside $3 billion in loss reserves for LDC debt, setting off a round of reserving among U.S. money center banks and large U.K. banks. Another round of reserving took place in late 1987 and early 1988, when U.S. regional banks initiated a large buildup of reserves, improvement of their equity capital ratios, and aggressive debt reduction strategies.

The most recent round of provisioning took place in September 1989, when U.S. money center banks strengthened reserves for LDC debt to an average level of 58 percent. Prior to this round, U.S. money center banks' loan loss reserves averaged only 25–30 percent—an insufficient level to embark on the size of debt reduction projected under the Brady Plan. It remains unclear whether the recent reserving was driven more by the necessity to position balance sheets for the Brady Plan, or by some banks' desire to wash their hands of LDC debt and the anticipated coerced Brady Plan deals. Likely, some of both factors has been at work. In the case of Morgan Guaranty, its decision to increase reserves to fully 100 percent of its LDC loans appears to indicate the latter strategy.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In early 1987, the weighted average price of LDC debt on the secondary market was just under 70 percent, indicating that the money center banks' provisioning to 25–30 percent was nearly adequate at the time. The continued price decline throughout 1987 and 1988, however, suggested the need for larger provisions. Still, some U.S. regulators in early 1989 maintained that U.S. banks were well positioned for any specter of loss. L. William Seidman, chairman of the Federal Deposit Insurance Corporation, told the House Banking Committee:

> Since 1982, the nine money-center banks have been successful in building their primary capital to a level which would allow them to withstand any likely event in the LDC arena. . . . [They] would continue to be solvent even if they wrote down to current secondary market levels all their exposures to the six major LDC countries. *Moreover, even in what surely could be considered a worst-case scenario, each of the nine money-center banks could write off 100 percent of their outstanding loans to these six countries and, on an after-tax basis, each of these banks would remain solvent* [emphasis in prepared testimony; Seidman 1989].

Two months after Seidman delivered this judgment that U.S. banks could withstand projected losses on LDC debt, U.S. Treasury Secretary Brady announced his new policy to, essentially, force them to do so. Yet, if there were any doubt in early 1989 that U.S. money center banks could go it alone with loan loss reserves of 25–30 percent, this doubt disappeared with the September 1989 reserving to an average of 58 percent. Given that LDC debt was selling at a weighted average of about 30 percent of face value in early 1989 (prior to announcement of the Brady Plan), the major U.S. banks are now close to being fully provisioned against the market's estimation of projected loss. Little doubt now remains that U.S. money center banks can go it alone. The still rather vulnerable balance sheets of Bank of America and Manufacturers Hanover could be aided by an immediate regulatory revision to allow for the full tax deductibility of loan loss reserves.

Brady Plan Only Stifles Private Debt-Reduction Activity

The new debt strategy that Brady announced on March 10, 1989, while initially a bare skeleton of a policy, appeared to debtors and their creditor banks to be the big public bailout for which they had waited, and lobbied, for several years. Indeed, investors recognized that Western taxpayers were about to underwrite the outcome of future debt-reduction deals. The market value of LDC debt began to rise immediately—only to fall again in late 1989 as the very high

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

early expectations for the plan subsided (see Table 5). Initially, the popular press was so swept up in reporting Brady's "bold new direction" that it was weeks before the fog cleared and certain realities settled in. Only modest sums of public money would be available from the World Bank, IMF, and Japan to back the initiative. This supply would produce only an estimated $6 billion in annual debt reduction over three years.

In the meantime, what happened to private sector debt conversion activity during 1989? Brazil suspended its debt-equity program on January 15, 1989. Argentina held no debt-equity auctions in the first half of the year. In the Philippines, the volume of debt conversions remained small. As Table 2 shows, seven countries carried out debt conversions for the first time in 1988; yet nine nations ceased conversion activity in 1989. Ecuador, for example, accounted for $258 million in debt swaps in 1988, but suspended all activity in early 1989. As one Ecuadoran official explained: "The Central Bank has said to hold off until after the [potential of the] Brady debt policy firms up" (Embassy of Ecuador 1989). Total debt conversion dropped off markedly, from $22.4 billion in 1988 to $13.8 billion in 1989 (see Table 2).

Conclusion

The present LDC debt crisis has much in common with previous episodes of sovereign debt default. Following a centuries-old tradi-

TABLE 5

SECONDARY MARKET PRICES OF DEVELOPING COUNTRY DEBT: THE TEMPORARY BRADY REBOUND

	Average Prices of Foreign Debt (in Cents per US$ of Face Value)			
	1/5/89	3/1/89[a]	7/17/89[b]	12/8/89
Argentina	22	18	19	13
Brazil	41	28	32	23
Chile	59	56	65	61
Mexico	42	34	44	37
Philippines	48	38	54	46
Poland	35	32	39	21
Venezuela	39	27	40	35
Yugoslavia	45	44	54	51

[a]Prior to Secretary Brady's March 10, 1989, announcement of his debt plan.
[b]Post–Brady Plan announcement; parties negotiating its first package (Mexico).
SOURCE: Salomon Brothers Inc. (January 5 to July 17) and Merrill Lynch (December 8).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tion, LDC debts to private creditors are being periodically renegotiated on ever-easier terms and stretched out across decades. Under the Brady Plan's debt-reduction agreements for Mexico and Venezuela, for example, transform large portions of these nations' medium-term debts into bonds with 30-year maturities (with taxpayer-financed interest guarantees).

Prior to Treasury Secretary Brady's 1989 initiation of a plan for officially assisted debt reduction, innovative debtors and private creditors were chipping away at ever-larger pieces of the debt by facing reality and making the best of a bad situation. Announcement of the Brady Plan served only to choke off much of the private sector's debt-reduction activity. But for the ever-larger, annual, new loan commitments from the World Bank and IMF, which serve to fill debtor nations' financing gaps, LDC debtors and their creditor banks would resolve a much larger share of the debt overhang through privately concluded debt-reduction opportunities, such as debt-equity swaps and straight debt buy-backs.

The eight-year-old U.S. strategy of papering over de facto LDC debt defaults with ever-larger levels of official lending has been laying the foundation for an unserviceable LDC debt burden to the World Bank and IMF. The Brady Plan adds only a new wrinkle to this slow transfer of private banks' LDC debt exposure to the World Bank and IMF (and the U.S. and other Western nation taxpayers who fund these institutions). Most U.S. banks are now very well provisioned against their LDC loans. U.S. policy should be to step out of the way and allow debtor nations and their creditor banks to drive their own hard bargains over these claims. The only consolation of the misguided Brady Plan is the likelihood that today's debtor fatigue over servicing private bank debts will become tomorrow's debtor fatigue with the growing multilateral institution debts. This extension of debts is likely to seriously jeopardize the financial integrity, and thus the staying power, of these two most pernicious of institutions of central planning—the World Bank and IMF.

References

Barth, James R., and Keleher, Robert E. "Is There a Role for an International Lender of Last Resort?" *The World of Banking* (November–December 1984): 25–27.

Bergsman, Joel, and Edisis, Wayne. "Debt-Equity Swaps and Foreign Direct Investment in Latin America." IFC Discussion Paper no. 2. Washington, D.C.: The World Bank, 1988.

Brooke, James. "Peru Trying to Shift Focus of Trade to Pacific." *New York Times*, 19 July 1989: p. 10.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Burns, Arthur F. "Statement on the Condition of the Banking System before the Senate Committee on Banking, Housing, and Urban Affairs." *Federal Reserve Bulletin* (March 1977): 238–46.

Camdessus, Michel. Transcript of a 21 September 1989 press conference. Washington, D.C.: The International Monetary Fund, 1989. Mimeo.

Durr, Barbara. "Chilean Debt Swaps Soar in First Half." *Financial Times* (London), 20 July 1989.

Embassy of Ecuador. Conversation with an economic attaché, Washington, D.C., 13 July 1989.

Eichengreen, Barry, and Portes, Richard. "Dealing with Debt: The 1930s and the 1980s." In *Dealing with the Debt Crisis*, pp. 69–86. Edited by Ishrat Husain and Ishac Diwan. Washington, D.C.: The World Bank, 1989.

Hanke, Steve H. "The Anatomy of a Successful Debt Swap." In *Privatization and Development*, pp. 161–68. Edited by Steve H. Hanke. San Francisco: Institute for Contemporary Studies Press, 1987.

Jorgensen, Erika, and Sachs, Jeffrey. "Default and Renegotiation of Latin American Foreign Bonds in the Interwar Period." In *The International Debt Crisis in Historical Perspective*, pp. 48–85. Edited by Barry Eichengreen and Peter Lindert. Cambridge, Mass.: MIT Press, 1989.

Lewis, Clifford M. "When Countries Go Broke: Debt through the Ages." *The National Interest* (Winter 1986/87): 46–52.

Marichal, Carlos. *A Century of Debt Crises in Latin America: From Independence to the Great Depression, 1820–1930*. Princeton: N.J.: Princeton University Press, 1989.

Mead, Gary. "Mistrust Fans the Flames of Troubled Privatization." *Financial Times* (London), 22 March 1989.

Reuters. Dispatch from Santiago, 11 January 1989.

Sachs, Jeffrey. "LDC Debt in the 1980s: Risk and Reforms." In *Crises in the Economic and Financial Structure*. Edited by Paul Wachtel. Lexington, Mass.: Heath, 1982.

Schuker, Stephen A. *American "Reparations" to Germany, 1919–33: Implications for the Third-World Debt Crisis*. Princeton Studies in International Finance no. 61. International Finance Section, Department of Economics, Princeton University, July 1988.

Schwartz, Anna J. "International Debts: What's Fact and What's Fiction." Address to the Western Economic Association, 2 July 1988. *Economic Inquiry* 27 (January 1989): 1–19.

Seidman, L. William. Testimony before the Committee on Banking, Finance, and Urban Affairs, U.S. House of Representatives, 5 January 1989.

Tammen, Melanie S. "Energizing Third World Economies: The Role of Debt-Equity Swaps," Backgrounder no. 736. Washington, D.C.: The Heritage Foundation, 1989.

Volcker, Paul A. "The Recycling Problem Revisited." *Challenge* (July–August 1980): 3–14.

Wallich, Henry C. "The Future of Latin American Dollar Bonds." *American Economic Review* 33 (June 1943): 321–35.

Wertman, Patricia. "The Baker Plan: A Remedy for the International Debt Crisis?" CRS Issue Brief, Congressional Research Service, updated 23 April 1990.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

World Bank. *World Debt Tables 1989–90*, Washington, D.C.: The World Bank, 1989b.
World Bank. *Annual Report 1989*. Washington, D.C.: The World Bank, 1989a.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Monetary Equilibrium and the Productivity Norm of Price-Level Policy

George A. Selgin

> To a simple fellow like myself it seems that the lower prices which increased production makes possible would benefit everybody, but I recognize there must be a flaw in my thinking, for increased productivity has not brought—and does not seem likely to bring—lower prices. Presumably there is some good reason for this. Will someone explain?[1]

Introduction

Now that the Phillips curve has disappeared, leaving an "empty place where it used to be" (Leijonhufvud 1981, p. 276), economists must come face to face with the problem of deciding how the price level ought to behave. They can no longer treat price-level policy as incidental to employment policy. Yet, rather than becoming an object of economic controversy, the place left vacant by the Phillips curve has become the exclusive, if somewhat barren, grazing ground of advocates of a stable consumer price level. These advocates appear to be winning the macroeconomic policy battle by default. The only challenge now facing them seems to be that of implementing price-level stabilization by means of a strict and unambiguous policy mandate.

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Assistant Professor of Economics at the University of Georgia. He wishes to thank Don Boudreaux, Richard Ebeling, David Fand, Lawrence H. White, Bill Woolsey, and Leland Yeager for their suggestions and criticisms. Research for this paper was supported by an F. Leroy Hill Fellowship from the Institute for Humane Studies. An earlier version of this paper was presented at the Ludwig von Mises Lecture Series at Hillsdale College, April 22–24, 1990.

[1]A former Archbishop of Wales, in a letter to the London *Times*, as quoted in Robertson (1963, pp. 11–12n).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Robert Black (1986, p. 790) for example, argues that price stability should be "the preeminent and perhaps even the unique goal of monetary policy." He adds that it should be enforced in an "automatic or quasi-automatic way" (p. 793). Others who have held similar views include Reynolds (1982, pp. 37–41); Barro (1986); Hetzel (1985); Meltzer (1986); Hall (1982, 1984a, 1984b); and Yeager (1986a). A strict policy of price-level stabilization is also supported by several well-known policymakers, including Jack Kemp, Richard Rahn of the U.S. Chamber of Commerce, and some members of the Board of Governors of the Federal Reserve System.[2] A recent Joint Resolution (H. R. J. Res. 409) introduced by the House Subcommittee on Domestic Monetary policy would, if adopted, require the Fed to achieve a stable price level ("zero inflation") within five years.

Although price-level stabilization is the only widely endorsed price-level policy today, there was a time, just over half a century ago, when prominent economists from numerous schools favored a different approach—the "productivity norm" of price-level behavior. Under this approach, the consumer price level is allowed to vary inversely with changes in unit real costs of production. In theory, the productivity norm is equivalent to stabilization of a price index of factors of production; in practice, it is roughly equivalent to the stabilization of per capita nominal income.

This paper offers a highly preliminary reconsideration of the case for a productivity norm of price-level policy as against price-level stabilization. In so doing, the paper also revives and expands upon some forgotten early criticisms of price-level stabilization, and it shows the shortcomings and potential dangers inherent in proposals for the strict enforcement of a constant price level.[3] Because the paper is meant to compare the productivity norm to the alternative of price stabilization, it should not be construed as an attempt to demonstrate that the productivity norm is an optimal or first-best policy.

The Rationale of Price-Level Stabilization

The alleged benefits of price-level stability—generally taken to mean stability of a consumer's price index—include avoidance of debtor-creditor injustice and avoidance of macroeconomic disequi-

[2]For example, Messrs. Johnson and Angell.

[3]This paper is only incidentally concerned, however, with problems of *implementing* various price-level policies (e.g., index-number and time-lag problems). These problems are well recognized and have been discussed recently elsewhere (e.g., in Garner 1985).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

librium. Debtor-creditor injustice is caused by unexpected changes in the value of long-term debts. Falling prices increase the burden of indebtedness, conferring a windfall gain on creditors, whereas rising prices do the opposite. Price-level stabilization prevents this injustice. However, because indexation schemes such as the tabular standard also avoid changes in the real value of debt, the principal advantage of a stable price level must be sought in its ability to combat short-run macroeconomic fluctuations.[4]

Leland Yeager has eloquently argued for the macroeconomic benefits of a stable price level. According to Yeager (1986b, p. 370), macroeconomic fluctuations—which can be taken to refer to fluctuations of employment and output around their "full information" or "natural" levels—are caused by monetary disequilibrium, that is, by "a discrepancy between actual and desired holdings of money at the prevailing price level." The occurrence of monetary disequilibrium implies (in the absence of instantaneous or even anticipatory price adjustments) a violation of Say's Law (though Yeager himself does not use this term): An excess demand for money implies a deficient effective demand for goods with concomitant windfall losses to producers; an excess supply of money implies an excessive effective demand for goods with concomitant windfall profits to producers. Because an excess supply of money leads to rising prices and a deficient supply leads to falling prices, general price changes can be viewed as "symptoms or consequences" of monetary disequilibrium (Yeager 1986b, p. 373). It follows that macroeconomic fluctuations will be avoided or reduced by a policy that adjusts the nominal money stock in such a way as to keep the price level stable.

Although it ultimately rests on a quantity-theoretic foundation, Yeager's macroeconomic defense of price-level stabilization contradicts simpler versions of the quantity theory in a crucial respect: It rejects the view that changes in the money supply or in its velocity of circulation lead to instantaneous, uniform, and costless adjustments in all prices. Were such a simple interpretation of the quantity theory valid, monetary disequilibrium could never exist for more than an instant, and there would be no macroeconomic reason for advocating any particular money supply or price-level policy. Rather than accepting this view, Yeager and like-minded proponents of price-level stabilization argue that general price adjustments "do not and cannot occur promptly and completely enough to absorb the

[4]Irving Fisher remarked (1925, p. 261) that, in practice, a tabular standard "would never accomplish more than a small fraction" of what price-level stabilization could achieve.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

entire impact of [a] monetary change and so avoid quantity changes" (Yeager 1986b, p. 373).

A number of reasons account for the sluggishness of general price changes. These include the presence of long-term contracts not subject to indexation and to other psychological sources of price rigidities such as implicit contracts and money illusion. Perhaps the most fundamental reason, though, has to do—according to Yeager (1986b, p. 377)—with the "public-good" nature of general price adjustments. This stems from money's role as a general medium of exchange that, "unlike other goods, lacks a price and a market of its own":

> No specific "money market" exists on which people acquire and dispose of money, nor does money have any specific price that straightforwardly comes under pressure to clear its (nonexistent) market. Money's value (strictly, the reciprocal of its value) is the average of individual prices and wages determined on myriads of distinct though interconnecting markets for individual goods and services. Adjustment of money's value has to occur through supply and demand changes on these individual markets.

The consequence is a diffusion of the impact of monetary disequilibrium across various markets, where each affected transactor regards the value of money "as set beyond his control, except to the utterly trivial extent that the price he may be able to set on his own product arithmetically affects money's average purchasing power" (Yeager 1986b, p. 392). Optimal adjustments in individual prices do not take place because their social value may exceed their perceived value to the persons who have to make them. This outcome is all the more likely given that particular price adjustments, rather than being costless as they are often portrayed, frequently involve lump-sum or "menu" costs—of printing, labeling, and negotiations. In consequence, rather than being achieved automatically following a monetary disturbance, a market-clearing general price level has to be "groped towards" by means of a "decentralized, piecemeal, sequential, trial and error" process (Yeager 1986b, p. 375).

Because price adjustments may be slow, they are also likely to be uneven—a result, in part, of the differing degrees of sluggishness of different prices. It is generally assumed that input prices adjust more slowly than product prices. This lag implies that excess demand for money will be a cause of painful short-run losses, whereas excess supply will lead to profit inflation.

A further cause of unevenness of price adjustments is the monetary "transmission mechanism" by which monetary disequilibrium makes its presence felt, not in all markets at once, but first in particular markets from which it slowly spreads to the rest of the economy

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

(Yeager 1990). Thus, a disequilibrium increase in the money supply on the basis of open-market purchases will first raise the value of government securities and then will affect general interest rates through an increase in the volume of commercial bank loans. From there, the monetary expansion will raise the demand for capital goods and increase wages. Only afterward will it lead to a more general increase in the prices of commodities. And this increase will occur even if prices all along the way are fairly flexible.[5]

Such imperfect price adjustments in response to monetary disequilibrium cannot fail to involve many temporary relative price effects that, by introducing "noise" into price signals, "degrade the information conveyed by individual prices" (Warburton [1946] 1951, p. 374) and provoke unwarranted changes in real activity. A shortage of money will lead to deflation, with reduced sales and production cutbacks in certain sectors of the economy leading to reduced demand for the products of other sectors and finally to general unemployment. An excess supply of money, on the other hand, causes inflation that, because it does not merely imply a uniform increase in prices, can also involve substantial malinvestment of resources.[6]

Responding to the potential dangers from imperfect adjustment of general prices, proposals for stabilizing the price level aim at *minimizing the burden placed on the price system* by maladjustments in the money supply. Adjusting the nominal quantity of money to keep the price level constant in the face of changing demands for real money balances is supposed to achieve this goal in two ways: first by reducing the overall requirement for permanent money-price changes, and second by reducing the extent of temporary, though ultimately unnecessary, relative price changes involved in the monetary transmission mechanism. These changes include disequilibrium movements in interest rates. In the absence of appropriate adjustments of the nominal money stock, both types of price changes must occur to some extent, and each will be a cause of disturbances to real activity. Only the permanent price changes will disturb real activity (apart from "menu-cost" effects) insofar as they fail to occur com-

[5]See also Warburton ([1946] 1951, pp. 298–99). Many contemporary monetarists are skeptical concerning the importance of "first-round effects" of changes in the quantity of money. See the discussion in Friedman and Schwartz (1982, pp. 29–31). Of course, to insist on the importance of such effects is not necessarily to agree with the Keynesian view that long-run effects can be ignored. The position of price-level stabilizationists—and also of proponents of the productivity norm (discussed below)—is essentially a Wicksellian one, which acknowledges both short-run and long-run effects while recognizing the difference between them.

[6]See Leijonhufvud (1984).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

pletely and uniformly, whereas relative price changes will disturb real activity to the extent that they occur at all.

Implicit in these arguments for price-level stabilization is the assumption that changes in real money demand or nomimal money supply, and consequent needs for general price adjustments, cannot be perfectly anticipated by economic agents. Although long-term tendencies in the movement of the equilibrium price level may come to be anticipated, short-run disturbances are likely to be completely unexpected and, hence, unrecognized for what they are. Knowledge of the pattern or policy of nominal money-supply changes (assuming this knowledge can be had) is not sufficient to avoid surprises; there may also be unexpected changes in real money demand for which scant public information is available.[7] This predicament brings to bear two further arguments for stabilizing the price level. One is that such stabilization reduces the uncertainty encountered by economic agents, allowing them better to capture potential gains from long-term contracts and production processes. The other is that it puts the monetary authorities on a tight leash by committing them to an unambiguous rule, violations of which are easily detected. The last argument is, however, more prominent and valid today than it was earlier in this century when monetary authorities were disciplined by the gold standard. In that context, price-level stabilization represented, at best, the substitution of one kind of monetary rule for another; at worst, it was a stepping stone from reliance upon a rule to reliance upon unrestrained authority.

The Productivity Norm

The Productivity Norm in the History of Thought

The productivity norm had many proponents before the ascendancy of Keynesian thought.[8] Perhaps its earliest champion was Samuel Bailey in his *Money and Its Vicissitudes in Value* (1837). Later British economists who at one time or another defended the productivity norm included Marshall, Edgeworth, Giffen, Hawtrey, Pigou, and Robertson.[9] In Sweden the norm was defended by David Davidson in a protracted debate with Wicksell (who in the end partially acquiesced), and also by Lindahl and Myrdal. Elsewhere in Europe the norm was embraced by German, Austrian, and Dutch writers of

[7]This fundamental point is overlooked by Grossman (1986) in his critical response to Yeager (1986b).

[8]Selgin (1988) treats the history of the productivity norm in some detail.

[9]Among Robertson's better-known students, Harry Johnson (1972, p. 29) also endorsed the productivity norm.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

the neutral money school, including Roepke, Mises, Hayek, Haberler, Machlup, N. G. Pierson, and J. G. Koopmans. New Zealand economist Allen G. B. Fisher (1935) defended the productivity norm at length in the *American Economic Review*. Finally, in the United States the norm was endorsed by Taussig, Laughlin, and Simon Newcomb in the 1890s and by John Williams in the 1930s. Some American champions of a stable price level, including Mints (1950, pp. 132–34) and Warburton ([1946] 1951, p. 308n), also conceded that price-level stability was not necessarily superior to the productivity norm.

In short, by the 1950s the productivity norm had received serious attention from economists of most schools. There was even a period—the first half of the 1930s—when it seemed to rival price-level stabilization as an ideal for monetary policy. Its popularity was short-lived, however, as it and all other prescriptions for macroeconomic stability were eclipsed by the views contained in Keynes's *General Theory*. The (perhaps unintended) consequence of Keynes's contribution was to detract attention from price-level policy altogether. What mattered was the achievement of full employment, regardless of what this required in the way of movements in the price level. In the end, this approach proved to be a recipe for inflation, which eventually drew economists' attention once more to the question of price-level policy. Only by then—in the 1970s—the productivity norm had fallen by the wayside, and price-level stabilization emerged by default as the sole, popular option for price-level policy.

The Rationale of the Productivity Norm in Formal Theory

The productivity norm rests upon the same tenets that underlie the norm of price-level stability. Both norms take for granted the desirability of a monetary policy that will combat monetary disequilibrium, while rejecting attempts to employ monetary policy to divert the economy from its natural or full-information levels of employment and output. The two norms also hold in common the assumption that the public's expectations may be less than fully correct in that individuals may fail to anticipate fully changes in income, real output, or the price level. More particularly, the norm of price-level stability implicitly assumes that individuals expect the price level to be *stable*—or at least that it is easiest for the public to form correct forecasts of price-level movements when such movements are altogether avoided.

The argument for the productivity norm, as opposed to a stable price-level norm, is that even in a situation where the price level has been kept stable for some time—say from t_o to t_n, so that the public

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

is firmly convinced that the price level will again remain unchanged at t_{n+1}—deviations from the natural rate of output will be smaller under a policy that allows the public to be surprised by a difference between p_{n+1} and p_n, provided the difference reflects a change in aggregate productivity. In this case an unanticipated change in the price level serves to compensate for an unanticipated change in real productivity. A formal demonstration of this is offered by Bean (1983), who models labor-market disturbances in an economy where money wages are set one period in advance and where there is uncertainty as regards both the price level and productivity. The goal of policy is to minimize the difference of output from its full-information level. Output obeys the expression:

$$y_t - y_t^* = \beta[(p_t - {}_{t-1}p_t) + \phi(u_t - {}_{t-1}u_t)]$$

with $(0<\beta<1)$, where y_t and y_t^* are the logs of the actual and expected values of the price level and productivity per worker (${}_{t-1}u_t + {}_{t-1}p_t$ thus equals the nominal wage).

If the supply of labor is inelastic with respect to changes in productivity, then $\phi = 1$ and changes in the price level should be fully proportionate to opposite changes in output. Such a policy is equivalent to one of stabilizing money income—$p_t + y_t$.

If $\phi < 1$ (i.e., if the supply of labor is elastic), the price level should adjust *less* than in proportion with changes in output, to allow for changes in the size of the full-information labor force.

In general, the price level should vary so as to stabilize *money income* per laborer—$p_t + y_t - l_t^*$—where l_t^* is the size of the full-information labor force. This condition is equivalent to saying that the price level should reflect changes in productivity: A *negative* productivity shock should be offset by a *positive* price-level shock, and a *positive* productivity shock should be offset by a *negative* price-level shock. A policy of stabilizing some measure of *per capita* money income ($p_t + u_t$) represents a practical approximation of this theoretical rule.

If the size of the full-information labor force is unchanging and if the demand for real money balances is unit-elastic with respect to changes in real income (and does not increase or decrease owing to causes not related to any change in real income), then adherence to the productivity norm will require that the nominal quantity of money be held *constant*. Unless otherwise stated, this case is the one considered in the arguments to follow. If, however, the demand for money is elastic with respect to changes in real income, then (other things being equal) an *increase* in productivity will require an *increase* in the nominal quantity of money, and a *decline* in produc-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

tivity will require a *reduction* in the nominal quantity of money *to prevent prices from falling or rising more than in proportion to the change in productivity*. In contrast, if the demand for money is inelastic, relative to changes in real income, a productivity norm policy will require that changes in productivity be accompanied by opposite changes in the nominal quantity of money.

Of course, changes in productivity need not always come as a surprise to economic agents. A secular increase in productivity, for example, may be perfectly anticipated in principle. But then the price-level trend (accompanying a productivity-norm policy) would also be perfectly anticipated and would, therefore, be no less desirable than any other fully anticipated price-level trend. Moreover, as will be argued below, a productivity-norm-based trend in prices is likely to be more consistent with the aim of allowing individual, *relative* prices to move in response to changes in productivity where such changes do not occur uniformly and predictably in all industries but are at any moment greater in certain industries than in others.

We now turn to consider how the above arguments hold up against two particular and intuitively appealing arguments for price-level stabilization: (1) that price-stability is needed to preserve debtor-creditor equity, and (2) that it is desirable for avoiding difficulties connected with sluggish price adjustment.

The Productivity-Norm and Debtor-Creditor Equity

Consider first the matter of debtor-creditor equity, where "debtors" include all persons who have committed themselves to making fixed-money payments in the future, and "creditors" include all persons who have agreed to receive these fixed-money payments. In a stationary economy where productivity is constant, it is generally agreed that debtors will suffer unjustified losses if the price level falls unexpectedly, and that creditors will realize unjustified gains. If, on the other hand, the price level is held constant, neither debtors nor creditors will (on the whole) have any reason to regret their involvement in contracts fixed in money terms.

But if productivity is changing, a stable price level may no longer achieve this desirable result. Assume, for example, that the public expects both the price level and productivity to remain stable. Then, if the price level is kept constant in the face of unexpected improvements in productivity, readily adjusted money incomes—including profits, dividends, and some wage payments—will increase; their recipients will benefit exclusively from the improvements in real output. Creditors, on the other hand, will not be allowed to reap any gains from the same improvements. Although a constant price level

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

may fulfill their price-level expectations, creditors may still regret their involvement in fixed-money contracts, for they may rightly sense that, had they anticipated the widespread improvement in other persons' real (and, in this case, money) earnings, they could have successfully negotiated better terms. On the other hand, if the price level is allowed unexpectedly to fall to reflect improvements in productivity, creditors will automatically enjoy a share of the gain, while debtors will have no reason to complain: Although the real value of their obligations rises (along with everyone else's), so does their real income. The burden of nominal payments imposed upon them is, however, unchanged. The debtors' only cause for regret is their missed opportunity to enjoy—owing to creditors' lack of perfect foresight—an undeserved windfall at the creditors' expense; their loss, as Haberler (1931, p. 21) put it, is only *lucrum cessans* and not *damnum emergens*.

Some people have objected (e.g., Mints 1950, pp. 132ff, and Haberler 1931, pp. 15–16) that this argument rests entirely on the premise that creditors deserve a share of improvements in productivity, and that no scientific grounds can be given in support of the argument. This objection leads to the conclusion that considerations of equity alone cannot provide any basis for choosing between the productivity norm and a stable price level.

But this conclusion appears to hold only if *improvements* in productivity alone are considered. In his 1889 memorandum to the Committee to Investigate Variations in the Value of the Monetary Standard, Edgeworth (1925, p. 222) observed that those who plead for stabilizing the money value of nominal debts in times of increasing prosperity "might be embarrassed if the principal were extended to the case of declining prosperity." Indeed, if productivity is falling—as during a negative supply shock—the inequity of a price-level stabilization rule cannot easily be denied, for in this case to keep the price level from rising requires a *contraction* of all nonfixed money incomes. This contraction adds to the burden of payment borne by debtors, increasing the likelihood that some or many of them will be unable to meet their obligations. As Lindahl, the Swedish follower of Davidson, observed, a price-level policy that may encourage parties to engage in unfulfillable agreements cannot be judged as equitable in any reasonable sense of the term (cited in Caplan 1942, p. 210).[10] In such cases it is clear that the productivity norm, rather

[10]Or, as Robertson (quoting Shylock) put it in *Money* ([1922] 1928, p. 121): " 'I'll have my bond, speak not against my bond'—is that a plea which should be listened to from a debenture-holder or Trade Unionist in a country shivering for lack of fuel or impoverished by chronic warfare?"

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

than a norm of price stability, best allows debtors and creditors to accomplish their goals and to avoid inequity when relying upon contracts fixed in money terms.

The Productivity Norm and Price Adjustment

The question still remains whether the productivity norm is superior to price-level stabilization in preserving short-run macroeconomic equilibrium. It may be recalled that a major advantage claimed for price-level stabilization in this regard is its alleged ability to minimize the burden of general adjustments borne by the price system. Here again, however, the advantage is no longer present when productivity changes, for *both the extent of necessary "permanent" price changes and the extent of temporary, but ultimately unnecessary, price changes are likely to be greater under price-level stabilization than they would be under the productivity norm.*

Suppose, for example, that 1,000 final goods are produced using three distinct factors of production only. A technological improvement causes the output per period of good x, which formally had a price (included in the price index) of one dollar, to double. Assuming (1) a constant money supply and velocity of circulation of money; (2) that x has a unitary price elasticity of demand; and (3) that demand for goods other than x is independent of real purchases of x, holding nominal income unchanged (thus abstracting from the need for any "secondary" relative-price adjustments), the price of x will fall to 50 cents. This implies a slight decline in the price index. Prices of all other goods, including the three factors of production, remain unchanged. The new equilibrium price structure requires one price adjustment only and represents an application of the productivity norm.

Now suppose, instead, that the price level is to be held stable under identical circumstances. To accomplish this, the authorities must expand the supply of money to achieve a uniform, though very slight, increase in the prices of 999 goods and of the three factors of production. The sole exception is good x, the price of which must (as in the previous case) still be allowed to fall, only less than in proportion with the improvement in its rate of output. This approach alone serves to keep the price index stable while also allowing needed adjustments in *relative* prices.

It is possible to construct examples in which the burden of price adjustment (reckoned in terms of the required number of permanent price changes) under price-level stabilization is *less* than what would be required under the productivity norm. This would be true, for instance, if there were a uniform increase in productivity for *all* final

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

goods, and if the number of different factors of production were less than the number of final goods. But such cases are so exceptional that they may safely be ignored in practice.[11]

Admittedly, arguments such as those made here concerning the burden of price adjustments under various price-level policies are distressingly dependent upon artificial assumptions. One must admit that, in reality, any single relative-price adjustment can be expected to have secondary effects. These effects lead to an all-around adjustment in relative prices and leave no grounds for preferring any one policy as minimizing the total number of required price adjustments. Nevertheless, I have tried to show that, insofar as *any* case can be made (by appropriately stringent assumptions) for a particular price-level policy using the price-adjustment criterion, it is one that favors the productivity norm rather than a stable price level. To the extent that it requires more price adjustments than the productivity norm, price-level stabilization increases the odds of price adjustments being imperfectly accomplished. It, therefore, tends to promote more widespread, undesirable changes in quantities from their full-information levels.

Another difference between price adjustments made necessary by unaccommodated changes in productivity and those made necessary by changes in the flow of money income (as must occur if the price level is to be kept stable in the face of productivity changes) is that the former are brought about through a more direct stimulus than the latter. The stimulus provided by productivity changes to equilibrium price movements is either immediate, as in the case of prices of goods the rate of output of which is altered (where price changes are a direct response to shifting supply schedules), or is of the second order of mediacy, as when changes in output of one group of goods lead to changes in demand for other goods because of the non-unitary price elasticity of demand of goods in the first group.

In contrast, the effects of changes in the flow of money income on equilibrium prices tend to be indirect. These effects involve shifts in demand schedules through a whole series of markets (depending on the precise nature of the monetary transmission mechanism)

[11]It is regrettable that many macroeconomic models adopt a one-commodity framework, with labor as a sole, distinct factor of production. Such models represent precisely the kind of exceptional circumstance in which a desired pattern of relative prices can be established with the same number of absolute price adjustments (following a change in productivity) regardless of whether a price-level stability or productivity norm is adhered to. To see how different price-level policies may take a different toll on the price system and thereby provide different opportunities for error, one must refer to a multi-commodity framework.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

before relative prices and the distribution of demand reach their final, equilibrium levels. As Warburton ([1946] 1951, pp. 298–99) observed, "The first change occurs at the point where the additional money is introduced into or taken out of the economy and is expressed in an increased or decreased demand for the goods and services desired by the persons directly affected by the change in the quantity of money." Such monetary injection effects are another source of unnecessary and undesirable adjustments in quantities, which will be greater under a policy of price-level stabilization than under a productivity-norm policy. The greater the degree of price and wage-rate rigidity, the more extensive such undesirable quantity adjustments will be.

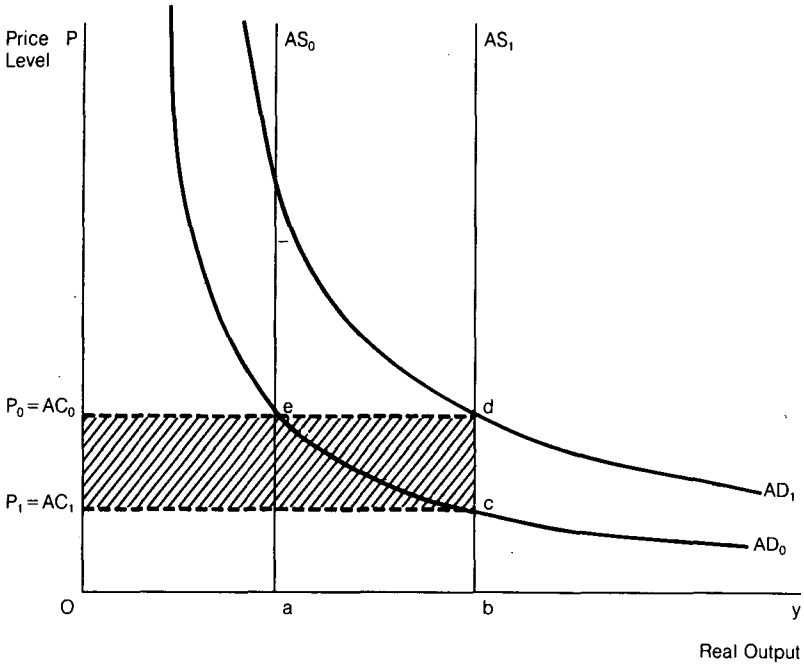
Besides being relatively direct and few in number, price adjustments in response to changes in productivity are also relatively easy and painless compared to price changes made necessary by changes in effective demand or in the flow of money income. This ease of responding to changes in productivity is still another reason why price adjustments are more likely to occur promptly. The reason is that productivity changes imply changes in unit costs of production. For a product with unitary price elasticity of demand, a change in the product's selling price equal to a change in its cost of production leaves the producer's revenues and profits unaffected. Such a change also does not place the producer under any pressure to negotiate new wage rates and salaries or to change the size of his work force. Figure 1 illustrates the case of a *general* increase in productivity caused, for example, by widespread technological innovation. Here, a doubling of real output per period, from "a" to "b," with a fixed quantity of factors of production and with an unchanged unit-elastic aggregate demand schedule, leads to a halving of the market-clearing price level. This result is consistent with an unchanged stock of money under the standard assumption that the demand for money is unitary elastic with respect to real income. If the elasticity of demand for money relative to real income is greater than or less than unity, the productivity norm requires expansion or contraction of the money supply to keep aggregate demand (MV, where M is the nominal money stock and V its velocity of circulation) from changing. Producers' aggregate revenues, formerly P_0Oae, afterwards P_1Obc, are not affected, and they suffer no losses. Because the reduction of prices required here is painless, there is no reason for producers to resist competitive pressures to undertake it.

A policy of price-level stabilization, in contrast, would require an expansion of money supply to shift aggregate demand to the right, from AD_0 to AD_1. This shift would make total revenues expand to

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FIGURE 1

A GENERAL INCREASE IN PRODUCTIVITY WITH UNIT-ELASTIC AGGREGATE DEMAND



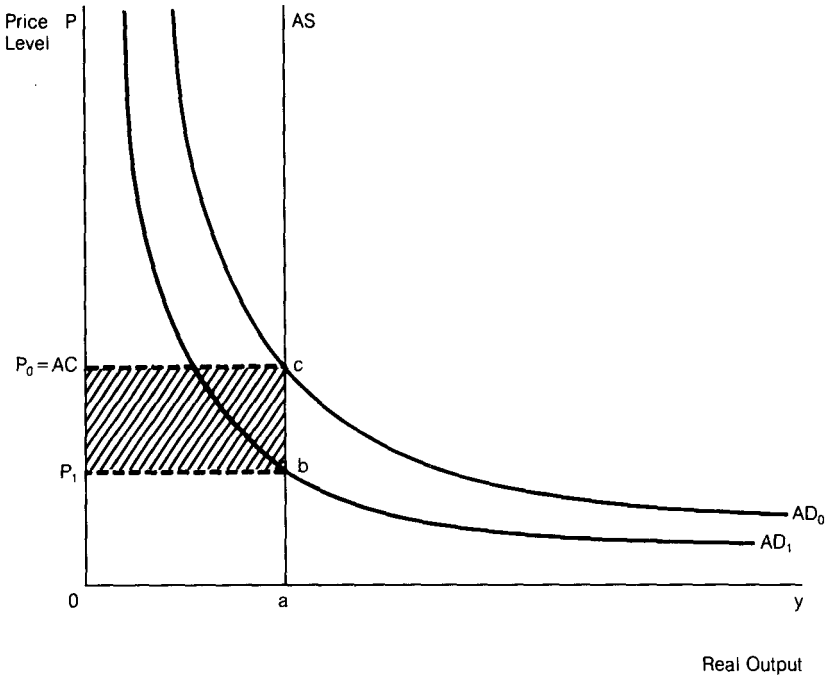
P_0Obd, causing profits to swell by the amount P_0P_1cd until factor costs adjust upward to eliminate the surplus. This upward adjustment of factor costs may be considerably more difficult and painful for producers to allow than the downward adjustment of prices required by the productivity norm. It is one thing to ask producers to pursue a pricing policy that serves merely to protect them from competition without affecting their profits; it is quite another to expect them to submit meekly to parting with extraordinary profits—even if only paper profits—once they have begun to enjoy them.

Figure 2 shows the contrasting case where the market-clearing price level falls by one-half because of an unanticipated decline in aggregate demand, from AD_0 to AD_1, with constant real output. Here, producers' aggregate revenues also shrink by one-half, from P_0Oac to P_1Oab. The loss on current output, represented by the shaded rectangle P_0P_1bc, will continue until nominal factor costs adjust downward. The process of price adjustment is evidently a painful one in this case. A policy of monetary expansion to keep aggregate

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

FIGURE 2

AN UNANTICIPATED DECLINE IN AGGREGATE DEMAND



demand and the price level at their original levels appears clearly preferable to one that would require them to fall.

By way of similar reasoning it can be seen that, although a rise in the price level resulting from expansion of aggregate demand should be avoided, a rise in prices of equal magnitude in response to a fall in productivity should be permitted. The only difference is that, in this case, the increase in prices in accordance with the productivity norm cannot be said to be truly painless, in that it represents a fall in real income and output. Nevertheless, to keep the price level stable under identical circumstances would be even *more* painful and discouraging to producers, because it would cause them to suffer temporary, nominal losses—incurring more than their fair share of the overall burden of reduced production—until laborers and other sellers of inputs are convinced that they must accept (painful) cost in wages and other factor prices.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In further support of these conclusions, we may note that many studies—both theoretical and empirical—of price rigidity incline to the view that product prices are rigid only because factor prices are rigid, and because product prices are often set according to a fixed percentage markup from costs.[12] Although this view accounts for the ill-adjustment of product prices in response to changes in effective demand, it does not predict any ill-adjustment in situations of changing productivity. In the latter case, unit costs of production are themselves changing, so that adjustments in product prices must take place, even as factor prices and the total outlay for factors stay the same, to preserve a constant markup. Implicit contracts between sellers of final goods and sellers of inputs are not necessarily violated when prices change in accordance with the productivity norm.[13]

Until now we have been assuming that changes in the market-clearing price level, whether caused by changes in productivity or by changers in aggregate demand, are completely unanticipated, which adds to the likelihood of needed price adjustments being for a time incomplete. But this assumption is less appropriate for price-level adjustments associated with changes in productivity, for the simple reason that changes in productivity are far less likely to be unanticipated *by price-setting agents in directly affected markets* than changes in aggregate demand. The reason, as Haberler (1931, p. 20) has observed, is that improvements in productivity are often (if not always) consciously aimed at by producers, who pursue the improvements precisely because they seek opportunities to profitably lower their prices compared to their rivals' prices. That downward price adjustments, caused by improvements in productivity, are often sought after by producers and are, therefore, anticipated is another reason for not regarding them as sluggish or incomplete.

All of the above considerations point to the fact that, under the productivity norm, aggregate producer revenues are more likely to keep in line with aggregate producer costs than under a stable price level if productivity is changing. If productivity is increasing, a stable price level requires expansion of money income, which (unless perfectly anticipated by all) must translate into a temporary, general swelling of firm profits. If productivity is falling, a stable price level

[12]This seems to have been Keynes's view in the *General Theory* (1936). See also Okun (1981, passim).

[13]On this point, see Okun (1981, pp. 169–70). Contrary to what Okun suggests, this is only one of several reasons accounting for the fact that "prices are much more responsive to changes in costs than to shifts in demand . . . even when demand is pulling in the opposite direction."

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

requires contraction of nominal income, which (unless anticipated) translates into a temporary, but general, fall in profits.

As Bertil Ohlin (1937, p. 321) once observed, keeping revenues "from getting out of balance" with costs (and especially wages) so as to prevent "abnormally large or low profits and profit expectations" is more crucial to macroeconomic order in the short run than stability of the price level as such. It is in this sense that the productivity norm, and not price-level stabilization, can be said to be truly consistent with the preservation of monetary equilibrium, where the latter requires continuous validation of Say's Law of markets. Only under the productivity norm will aggregate (effective) demand continue to be just adequate to buy the fruits of industry at prices covering their (money) cost of production, without causing that cost to alter over time except in response to growth in capital and population. In Wicksellian terms, the productivity norm manages, where stability of the price level fails, to keep interest rates at their natural levels.

The Productivity Norm and the Optimal Quantity of Money

Minimizing departures of relative prices and output from their full-information levels is only one possible objective of monetary policy. Another is to maximize consumers' welfare from money holdings. In general this calls for allowing prices of final goods to decline at a rate roughly equal to "the" real rate of interest (Friedman 1969). The productivity norm, by allowing the price level to decline secularly as productivity grows, comes closer to this ideal than price-level stabilization. Moreover, if the real rate of interest is equal to the rate of growth of real income (as has been roughly true historically and as is implied by standard models of optimal growth, e.g., Phelps 1961), the rate of deflation that maximizes consumers' welfare from money holdings will equal the rate of increase in per capita output, and will imply that factor prices are held constant. The productivity norm and the optimal quantity of money norm will then coincide.

The Productivity Norm in Practice

Though past upholders of the productivity norm agreed that it was tantamount to "stabilization of some average of the prices of the original factors of production" (Hayek [1933] 1984, p. 161), they differed in their views concerning how this stabilization could be achieved in practice.[14] It was generally admitted that a true index of

[14]The goal of stabilizing the "average price of factors of production" is not equivalent to one of stabilizing total money costs of production unless the quantity of productive inputs—of labor and capital—remains fixed: The productivity norm requires that prices fall to reflect *intensive*, but not extensive, improvements in output. Monetary expansion or contraction involving expansion or contraction of total money costs is required in connection with population growth and the accumulation of capital.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

prices of factors of production could not be constructed, because of the lack of market-price statistics for most factors of production apart from unskilled labor. But such an approach is by no means necessary or even appropriate: A productivity-norm policy can be implemented, or approximately implemented, by directly stabilizing some readily available measure of the flow of money payments or income such as nominal GNP or domestic final demand.[15]

Just how money income should behave if the productivity norm is to be enforced depends on the extent of changes in the quantity of factors of production—including both labor and capital. A common view emerged in the 1930s that money income per capita, rather than aggregate money income, should be stabilized. The stabilization of per capita income was meant to prevent growth in the labor force—assumed to be approximated by growth in population as a whole—from having a depressing effect on nominal wage rates. To this opinion Robertson (1957, p. 39) added that, if capital increases and, therefore, contributes a greater share to total production, then the part of money incomes imputable to rental payments should also increase. Thus, nominal income should grow at a rate sufficient to reflect changes in both the labor force and the capital stock. This would allow prices to fall in proportion to purely *intensive* improvements in output, but less than in proportion to any improvements of a partly extensive nature.

Viewed as a proposal for stabilizing nominal income, the productivity norm resembles a number of recent suggestions for targeting nominal GNP.[16] A major difference, however, is that many of these recent suggestions consider the targeting of nominal GNP as a means for achieving stability of the price level. Indeed, these suggestions have been criticized for being incompatible with price-level stability whenever supply shocks occur (e.g., Barro 1986, p. 26). The productivity norm, in contrast, does not value price-level or real-output stability per se; it seeks merely to avoid those price and output fluctuations that are likely to involve departures of economic activity from its full-information structure. This characteristic also distinguishes the productivity norm from proposals for an elastic price standard, such as those of Hall (1984b, 1986) and Taylor (1980).

[15]Haraf (1986) argues that targeting domestic final demand (DFD) (= GNP less changes in business inventories and net exports) is more advisable than targeting GNP. DFD may well be a superior measure of "consumers' outlay" or aggregate spending.

[16]See, for example, McCallum, (1984, 1985, 1987); Gordon (1985); Hall (1981, 1983); Haraf (1986); Tobin (1980, 1983); and Meade (1978). For critical discussions of some of these proposals, see Bean (1983) and Taylor (1985).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Although our main concern in this essay is not with the practical problems of implementing various price-level policies, it is appropriate to remark here that a nominal income target is no more difficult to implement than any price-index target. On the contrary, as Haraf (1986, p. 659) observes, an income target would probably lead to quicker reaction by the monetary authorities to velocity disturbances than a price-level target, without inviting inappropriate responses to real (productivity) disturbances.

Conclusion

Theoretical considerations and assumptions identical to those used by proponents of a stable price level lead one to favor the productivity norm over price-level stability and also over other price-level policies as a means for promoting general welfare. The norm would require the monetary authorities to stabilize per capita nominal GNP or domestic final demand or some other measure of per capita money payments, allowing for more rapid growth of money payments when the share of income attributable to capital is increasing.

Some features of a productivity norm—based on nominal income targeting—that should contribute to the norm's overall appeal include the following:

1. The fact that the productivity norm, like other popular policies, allows full monetary accommodation of changes in the velocity of money, with less reaction lag-time than a stable price-level norm and, hence, with reduced velocity-shock related variability of prices.
2. The long-run equivalence of a productivity norm to price-level stabilization under stationary output conditions.
3. The productivity norm's equivalence to a money growth-rate rule if velocity is stable.
4. The productivity norm's emphasis on stability of aggregate demand, which conforms well with natural rate theories of output and employment and with the rational-expectations view that macroeconomic instability is caused mainly by problems of "signal extraction."
5. The fact that the productivity norm is a more moderate policy than a policy of either price-level stabilization or full-employment—that is, the productivity norm avoids extremes of price- or employment-variability possible under these other policies.
6. The greater symmetry of the productivity norm compared with other price-level policies: Those who plead for monetary "accommodation" of adverse supply shocks do not also recom-

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

mend monetary contraction to counteract the real effects of positive supply shocks.
7. The relative ease of implementing the productivity norm.
8. The coincidence of the productivity norm with the optimum quantity of money norm when the real rate of interest is equal to the rate of growth of real output.
9. The ability of the productivity norm to provide protection from the monetary authorities' abuse of their powers equal to or greater than what could be achieved by other price-level norms.

A proposal for targeting nominal income (rather than any commodity price index) is neither novel nor necessarily controversial. Yet this particular form of the proposal may be controversial for the simple reason that it would allow the price level to *fall* in normal, progressive times. Though this approach was once widely endorsed by economists, it is now practically unheard of. Instead, mild inflation is taken to represent a normal and moderate condition, while price-level stability is regarded as a hard-line extreme.

It must be remembered, however, that policy itself is responsible for these views, because policy created the last four decades of inflation, which everyone now takes for granted. The policy in question was founded on the view that expansion of nominal (effective) demand can eliminate unemployment. We now know that this view is mistaken. Although unemployment may sometimes be due to a deficiency of effective demand, and although expansion of demand may temporarily reduce unemployment even when its ultimate cause is *not* deficient demand, the view that inflation as such reduces unemployment (i.e., that the long-run Phillips curve is negatively sloped) has been discredited. If policymakers are prepared to admit that inflation has no such benefit, then it is encumbent upon them to press beyond the simple analytics of a nonexistent "inflation-unemployment trade-off" in deciding how the price level ought to behave. Price-level stability is one answer; the productivity norm is another, better answer.

References

Bailey, Samuel. *Money and Its Vicissitudes in Value.* London: Effingham Wilson, 1837.

Barro, Robert J. "Rules vs. Discretion." In *Alternative Monetary Regimes,* pp. 16–30. Edited by Colin D. Campbell and William R. Dougan. Baltimore: Johns Hopkins University Press, 1986.

Bean, C. "Targeting Nominal Income: An Appraisal." *Economic Journal* 93 (December 1983): 806–19.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Black, Robert P. "A Proposal to Clarify the Fed's Policy Mandate." *Cato Journal* 5 (Winter 1986): 787–95.

Caplan, Benjamin. "The Wicksellian School—A Critical Study of the Development of Swedish Monetary Theory, 1898–1932." Ph.D. dissertation, University of Chicago, 1942.

Edgeworth, F. Y. *Papers Relating to Political Economy.* London: The Royal Economic Society, 1925.

Fisher, Allen G. B. "Does an Increase in Volume of Production Call for a Corresponding Increase in Volume of Money?" *American Economic Review* 25 (June 1935): 197–211.

Fisher, Irving. *Stabilizing the Dollar.* New York: Macmillan, 1925.

Friedman, Milton. "The Optimum Quantity of Money." In *The Optimum Quantity of Money and Other Essays,* chap. 1. Chicago: Aldine, 1969.

Friedmon, Milton, and Schwartz, Anna J. *Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices, and Interest Rates, 1867–1975.* Chicago: University of Chicago Press, 1982.

Garner, C. Alan. "Commodity Prices and Monetary Policy Reform." Federal Reserve Bank of Kansas City, *Economic Review* 70 (February 1985): 7–21.

Gordon, Robert J. "The Conduct of Domestic Monetary Policy." In *Monetary Policy in Our Times.* Edited by A. Ando, et al. Cambridge, Mass.: MIT Press, 1985.

Grossman, Herschel I. "Money, Real Activity, and Rationality." *Cato Journal* 6 (Fall 1986): 401–8.

Haberler, Gottfried. "The Different Meanings Attached to the Term 'Fluctuations in the Purchasing Power of Gold' and the Best Instrument or Instruments for Measuring such Fluctuation." Geneva: League of Nations, March 1931.

Hall, Robert E. "Lowering Inflation and Stimulating Economic Growth." In *Politics and the Oval Office: Toward Presidential Governance,* pp. 207–27. San Francisco: Institute for Contemporary Studies, 1981.

Hall, Robert E. "Explorations in the Gold Standard and Related Policies for Stabilizing the Dollar." In *Inflation.* Edited by Robert E. Hall. Chicago: University of Chicago Press, 1982.

Hall, Robert E. "Macroeconomic Policy under Structural Change." In *Industrial Change and Public Policy.* Kansas City: Federal Reserve Bank of Kansas City, 1983.

Hall, Robert E. "A Free-Market Policy to Stabilize the Purchasing Power of the Dollar." In *Money in Crisis.* Edited by Barry N. Siegel. San Francisco: Pacific Institute, 1984a.

Hall, Robert E. "Monetary Policy with an Elastic Price Standard." In *Price Stability and Public Policy,* pp. 137–59. Proceedings of a symposium sponsored by the Federal Reserve Bank of Kansas City, 1984b.

Hall, Robert E. "Optimal Monetary Institutions and Policy." In *Alternative Monetary Regimes,* pp. 229–39. Edited by Colin D. Campbell and William R. Dougan. Baltimore: Johns Hopkins University Press, 1986.

Haraf, William S. "Monetary Velocity and Monetary Rules." *Cato Journal* 6 (Fall 1986): 641–62.

Hayek, F. A. "On 'Neutral Money.' " 1933. Reprinted in *Money, Capital, and Fluctuations: Early Essays,* pp. 159–62. Chicago: University of Chicago Press, 1984.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Hetzel, Robert. "A Mandate of Price Stability for the Federal Reserve System." *Contemporary Policy Issues* (1985): 50–67.

Johnson, Harry G. *Inflation and the Monetarist Controversy*. Amsterdam: North Holland, 1972.

Keynes, J. M. *The General Theory of Employment, Interest, and Money*. London: Macmillan, 1936.

Leijonhufvud, Axel. *Information and Coordination*. New York: Oxford University Press, 1981.

Leijonhufvud, Axel. "Inflation and Economic Performance." In *Money in Crisis*, pp. 19–36. Edited by Barry N. Siegel. San Francisco: Pacific Institute, 1984.

McCallum, Bennett T. "Monetarist Rules in the Light of Recent Experience." *American Economic Review* 74 (May 1984): 388–91.

McCallum, Bennett T. "On the Consequences and Criticisms of Monetary Targeting." *Journal of Money, Credit, and Banking* 17 (November 1985): 570–97.

McCallum, Bennett T. "The Case for Rules in the Conduct of Monetary Policy: A Concrete Example." Federal Reserve Bank of Richmond, *Economic Review* 73 (October 1987): 10–18.

Meade, J. E. "The Meaning of Internal Balance." *Economic Journal* 88 (September 1978): 423–35.

Meltzer, Allan H. "Some Evidence on the Comparative Uncertainty Experienced under Different Monetary Regimes." In *Alternative Monetary Regimes*, pp. 122–53. Edited by Colin D. Campbell and William R. Dougan. Baltimore: Johns Hopkins University Press, 1986.

Mints, Lloyd W. *Monetary Policy for a Competitive Society*. New York: McGraw Hill, 1950.

Ohlin, Bertil. "Employment Stabilization and Price Stabilization." *Lessons of Monetary Experience*, pp. 318–28. Edited by Arthur D. Gayer. New York: Rinehart & Company, 1937.

Okun, Arthur. *Prices and Quantities: A Macroeconomic Analysis*. Washington: Brookings Institution, 1981.

Phelps, Edmond. "The Golden Rule of Accumulation." *American Economic Review* 53 (September 1961): 638–43.

Reynolds, Alan. "The Trouble with Monetarism." *Policy Review* 21 (Summer 1982): 19–41.

Robertson, Dennis H. *Money*. Cambridge: Cambridge University Press, 1922; 3rd rev. ed., 1928.

Robertson, Dennis H. *Lectures on Economic Principles*. London: Staples, 1957.

Robertson, Dennis H. "A Memorandum Submitted to the Canadian Royal Commission on Banking and Finance." Princeton University, *Essays in International Finance* (May 1963): 241–74.

Selgin, George A. "The Price Level, Productivity, and Macroeconomic Order." Unpublished manuscript, 1988.

Taylor, John B. "Output and Price Stability: An International Comparison." *Journal of Economic Dynamics and Control* 2 (February 1980): 109–32.

Taylor, John B. "What Would Nominal GNP Targeting Do to the Business Cycle?" *Carnegie-Rochester Conference Series on Public Policy* 22 (1985): 61–84.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Tobin, James. "Stabilization Policy after Ten Years." *Brookings Papers on Economic Activity* 1 (1980): 19–72.

Tobin, James. "Commentary." *Industrial Change and Public Policy*. Federal Reserve Bank of Kansas City, 1983.

Warburton, Clark. "The Misplaced Emphasis in Contemporary Business-Fluctuation Theory." 1946. Reprinted in American Economic Association, *Readings in Monetary Theory*, pp. 284–318. Homewood, Ill.: Richard D. Irwin, 1951.

Yeager, Leland B. "Price-Level Stability as the Goal of Monetary Reform." *Cato Journal* 5 (Winter 1986a): 821–24.

Yeager, Leland B. "The Significance of Monetary Disequilibrium." *Cato Journal* 6 (Fall 1986b): 369–99.

Yeager, Leland B. "Injection Effects and Monetary Intermediation." Unpublished manuscript, 1990.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

COMMUNICATIONS

WHY IS EDUCATION PUBLICLY PROVIDED? COMMENT ON LOTT

Richard B. Coffman

In a recent article in this journal, John R. Lott, Jr. (1987), argued that the best explanation for public provision of education is the need of politicians to indoctrinate the populace to accept the government as "fair" and "legitimate," in order to reduce the costs of political actions such as wealth transfers. Public choice economics, however, suggests at least three reasons why this hypothesis seems of doubtful validity for the United States.

First, the marginal value of indoctrination would seem slight where rational voter ignorance already has reduced voter opposition to wealth transfers to a low level. Voters are said to be rationally ignorant if their personal cost/benefit calculations tell them it is not rational to be informed about politics. The literature argues (a) that most voters face a high opportunity cost of diverting time and mental effort from their private affairs to consideration of public matters and (b) see little chance of influencing political outcomes by voting or otherwise participating in politics (Gwartney and Stroup 1989, pp. 94–95). Thus most voters remain rationally ignorant and pose little obstacle to special interest politics. If this is the case, then there would seem to be little payoff to further efforts to use indoctrination to reduce opposition to rent seeking.[1]

Second, indoctrination is an investment with a relatively long payback period, and as such will not be undertaken by politicians with short time horizons. The public choice literature provides consider-

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Associate Professor of Economics at the University of Idaho.

[1]In this paper the term "rent seeking" will be used to denote "actions by individuals and interest groups designed to restructure public policy in a manner that will either directly or indirectly redistribute more income to themselves" (Gwartney and Stroup 1989, p. 94).

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

able evidence that politicians take shortsighted actions because their time horizon is effectively limited to the next election (Gwartney and Stroup 1989, pp. 747–49). Using the educational system to indoctrinate citizens, though, would not pay off in the short run, since children do not vote, and some time would have to elapse before a significant proportion of the voting population would consist of those who had passed through the indoctrination pipeline.

Lott has addressed this issue in a more recent paper, suggesting four mechanisms that might make a long-term investment in indoctrination through schooling politically viable in a democracy (Lott 1990). First is the argument that "institutions that have the lowest costs of transferring wealth will produce the most political support and will be least likely to be challenged" (Lott 1990, p. 210). Of the several sources Lott cites for this argument, Gary Becker's 1976 paper seems most relevant. But Becker's model depends in part on the assumption that "voters perceive correctly the gains and losses from all policies" (Becker 1976, p. 247). Yet, as Becker notes, others have disagreed, arguing voters underestimate the costs of rent-seeking policies (Becker, p. 246). Further, if voters are rationally ignorant, they are systematically unaware of the costs of policies that have small effects on them personally. In this case, special interest voters will be the only ones expressing themselves on rent-seeking issues, and they will be likely to choose redistribution methods that maximize the gains to themselves, rather than those that minimize society's dead-weight loss of transferring wealth.

Lott's second argument is that "long-lived, strong political parties [might] punish politicians who fail to support investments in indoctrination" (Lott 1990, p. 211). However, he himself questions the importance of this argument since he cannot find empirical evidence to support it. It might also be argued that the long life of political parties does not necessarily imply continuity of policy since party control is repeatedly captured and lost by factions supporting different presidential candidates.

His third argument is that "constituencies that will receive these long-term benefits could support only politicians who vote to inculcate the desired values" (Lott 1990, p. 212). But this argument just moves the problem back from the level of the politicians to the constituents. Why should constituents use up political capital today to support a program with distant, vague, and uncertain future payoffs? Presumably indoctrination improves the general atmosphere for rent seeking without giving an advantage to any particular rent seeker. All future rent seekers will face competition that may dissipate their rents. Thus the returns from indoctrination are highly

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

uncertain. Politicians who promise future rent-seeking opportunities may well be less attractive to constituents than those who specialize in delivering even small current redistribution.

Finally, Lott suggests that politicians may support indoctrination for ideological reasons, even if it is not justified on pure vote-seeking grounds. Politicians may intrinsically value "education's moral indoctrination and/or the resulting future transfers" (Lott 1990, p. 212). Lott's reference to moral indoctrination is somewhat puzzling, since it is political indoctrination that is under discussion. Further, it is not clear what sort of ideology favors transfers per se. On the other hand, if political indoctrination is very similar to inculcating basic patriotism, then it is easy to understand how politicians might support it on "ideological" grounds. But then the whole argument has changed since they are no longer supporting it as a means to rent seeking. Finally, whether ideological support can be generated on the basis of specific patterns of future transfers would seem uncertain, since, as argued above, competition and the potential for rent dissipation make the patterns themselves uncertain.

It remains to discuss a third public choice objection to Lott's basic thesis. Public choice theory suggests that the value of indoctrination to politicians should be a function of the level of transfer activity allowed by the Constitution. Terry L. Anderson and Peter J. Hill have argued that the United States was a constitutional democracy prior to 1877, and turned into a majoritarian democracy by stages in the years up to World War I (Anderson and Hill 1980, pp. 54–68). Thus it is argued that the scope for rent seeking before this time was very limited. If this is correct, there would have been little payoff in the earlier time to investments in public education as a method of indoctrination.

However, the move to public schools seems to have started well before 1877. Public elementary education was well established by 1870, with 57 percent of all persons 5–18 years of age enrolled in public schools during 1869–70 (Dexter 1971, p. 164). Public high schools developed somewhat more slowly, but in cities larger than 25,000 population, 82 public high schools were established prior to 1870, and only 58 between 1870 and 1900 (Dexter 1971, p. 173). A somewhat different pattern is found for establishment of public high schools nationally, with only 349 established prior to 1870, 479 established in the 1870s, and 2,149 established between 1880 and 1900 (Dexter 1971, p. 173). In general, though, the data suggest the move to public provision of education was well under way before the time Anderson and Hill identify with the birth of a transfer society in the United States.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

In sum, these public choice arguments raise some doubts about Lott's hypothesized explanation for public provision of education in the United States.

References

Anderson, Terry L., and Hill, Peter J. *The Birth of a Transfer Society.* Stanford, Calif.: Hoover Institution Press, 1980.

Becker, Gary. "Comment on Peltzman." *Journal of Law and Economics* 19 (1976): 245–48.

Dexter, Edwin Grant. *A History of Education in the United States.* 1906. Reprint. New York: Lenox Hill (Burt Franklin), 1971.

Gwartney, James D., and Stroup, Richard L. *Economics: Private and Public Choice.* 5th ed. San Diego: Harcourt, Brace, and Jovanovich, 1989.

Lott, John R., Jr. "Why Is Education Publicly Provided? A Critical Survey." *Cato Journal* 7 (Fall 1987): 475–501.

Lott, John R., Jr. "An Explanation for Public Provision of Schooling: The Importance of Indoctrination." *Journal of Law and Economics* 33 (April 1990): 199–231.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Why Is Education Publicly Provided? Some Further Thoughts

John R. Lott, Jr.

In a series of papers, I have tried to provide a consistent explanation for the puzzle of why all the countries in the world have adopted public provision of schooling despite what appears to be its relatively high costs. My hypothesis proposed that public provision of education, like public provision of the media and even the use of coercion, is just another device that governments have to reduce the effective level of opposition arising from wealth transfers. People's views were assumed to be a function of the information that they receive. If governments can raise the costs of citizens receiving anti-transfer information and lower the costs of their receiving pro-transfer information, views more sympathetic to these transfers will be produced and the marginal opposition arising from an additional dollar of transfers will be reduced.

This simple hypothesis was shown to explain a variety of observable phenomena (Lott 1990c): (1) why higher government transfers and higher levels of totalitarianism are associated with increased expenditures on schooling, (2) how government expenditures on schooling vary in the same systematic way as government ownership of television and radio stations, (3) why exclusive territories are used for public schooling but not for other publicly provided goods, and (4) why increased opposition to government is associated with higher rents to educators.[1] The systematic nature of these features of public schooling across countries seems to argue against an approach where

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

The author is Visiting Assistant Professor at the John E. Anderson Graduate School of Management at the University of California at Los Angeles. He would like to thank Gertrud Fremling for her helpful comments.

[1]Recent work by Iannacone (1990) has sought to expand this discussion to help explain the creation of state-supported religions and some aspects of the institutional arrangement of religion.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

we rely on a different theory to explain each of these different features and why each individual country has adopted public provision.

Richard Coffman's paper contains three central points: (1) that free-rider problems produce voters who are largely unable to oppose governmental wealth transfers, (2) that politicians have an extremely short time horizon that does not extend past "the next election," and (3) that for the United States "some doubts" exist about the view that public provision of education can be explained as a method of lowering the costs of governmental wealth transfers. All three of these issues have already been discussed in my previous work as possible objections and I have tried to deal with them, though evidently not to Coffman's satisfaction.

The claim that voter opposition to transfers can essentially be ignored since voters obtain such a small return to voting relative to its costs seems both difficult to reconcile with the received view of support maximization and, if his discussion were the whole story, would make it difficult to explain why we observe as many people voting as we do (e.g., see Coursey and Roberts 1990). Even if the level of marginal opposition from additional transfers is low, it can still pay to reduce that opposition even further. Politicians maximize support by creating transfers until the marginal support generated by an additional dollar of transfer equals the marginal opposition that transfer creates. Reducing opposition, even from an already low initial level, allows for a higher level of transfers and increased support for a politician.

Even if the marginal opposition for a given level of transfers equaled zero, the most likely consequence seems to be higher transfers. It is unlikely that the special interests that Coffman refers to are so satiated that they are indifferent to whether or not they receive an additional dollar from the government. If the marginal support is indeed greater than the marginal opposition, transfers will rise and so too will the level of marginal opposition. At some point, the level of opposition will be sufficiently high that it will pay for some political entrepreneur to enter and propose reduced levels of transfers. Given the empirical evidence concerning the importance of the median voter or on the desire of politicians to maximize political support, it seems hard to believe that the politicians only have to look at the marginal support side of the equation.

Most of Coffman's discussion, however, focuses on the asserted inability of either politicians or their constituents to internalize the investments from indoctrination—that "their time horizon is effectively limited to the next election." This argument stems from a belief that it is only the threat of the most imminent reelection that

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

restrains how politicians vote, and that voters are unable to determine whether politicians are acting in their long-run interests. Yet, considerable empirical evidence indicates that members of congress in their last term of office, who no longer face the threat of reelection, do not change how they vote when they do vote (Lott 1987a; Van Beek 1990). Many find little or no evidence of politicians deviating from their constituents' interests (e.g., Peltzman 1984; Lott and Davis 1990; Bender 1990, 1991; and Goff and Grier 1990).[2] Why is it so difficult to believe that those constituents who benefit from these transfers care about the long-term effects of government actions (not just what happens between now and the next election) and, therefore, will reward those politicians who act in their interests?

Coffman's assertion that "presumably indoctrination improves the general atmosphere for rent seeking without confering any advantage to any particular rent seeker" is difficult to interpret. If investments in indoctrination result in a higher level of taxes, those groups whose level of transfers will rise with the level of taxes should anticipate this and act accordingly. The fact that competition will occur for these increased tax revenues does not obviate the claim that certain groups are better positioned to benefit from higher tax revenue than others. Nor is it obvious why the benefits of indoctrination cannot be

[2]While Coffman does mention that one of my papers (Lott 1990a) finds evidence consistent with political parties not effectively punishing politicians who reduce their attendance rates when they no longer face the threat of reelection, he fails to note that the paper dealt only with the case of the United States and that other democratic countries may have much stronger political parties or that the problem might not apply to more totalitarian ones. Coffman also claims that American democracy is relatively unstable and thus relatively unable to internalize the returns to long-term investments because "party control is repeatedly captured and lost by factions supporting different presidential candidates." First, the different groups that win control of political power may find it in their interests to use indoctrination to help accomplish their own ends. In India, for example, opposition parties have campaigned for abolishing the government's monopoly on television, but they have continually failed to abolish the monopoly once they have won office—instead preferring to use the government monopoly for their own benefit. Second, the fact that political power is so widely dispersed throughout the government and that there is such extremely high reelection rates in both houses of Congress implies that there is a great deal of continuity in political control over time. Third, even if different combinations of coalitions win the presidency, it is not clear that we are talking about anything other than very slight changes in the level of transfers if political markets are competitive and if the distribution of voters is unimodal. Most of the same groups will still receive net transfers though the amounts will vary somewhat at the margin. For example, the growth rate of government spending has not varied greatly across our last several presidents, and it does not seem as if there was a very dramatic change in the composition of spending. My (1990c) paper provides citations for my other work that deals with the question of which types of governments tend to be the most stable over time.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

targeted to aid certain groups relative to others (e.g., Lott 1987c, 1987d).

The third objection involves how my discussion relates to the United States. Historically, the desire to instill certain types of views in students has helped explain the use of exclusive territories in assigning students to public schools in the United States (Lott 1987b, 1990c). When one examines the United States both over time and across states, the level of education seems to be an important determinant of the level of governmental transfers (Peltzman 1980, pp. 270–76). The United States was also included in the cross-country evidence that indicates a very strong relationship in the 1970s between the level of transfers and the level of public expenditures on education (Lott 1990c). It is not difficult to find current examples of how public schooling affects the incentives of public educators to produce information and the effect that this has on the level of transfers (Lott 1990b). With respect to the claim that public schooling started before there were large-scale transfers in the United States, public schooling may have been one of the investments made in lowering the costs of creating those future transfers. For example, the high salaries paid to public teachers can create subtle biases as they produce information to defend public schooling and in the process create justifications that can be used for other publicly provided services (Lott 1990c).

Transfers can also take many forms other than the simple transfer of money. For example, one of the major motivations for the original introduction of public schooling in the United States was the desire to instill certain religious views (Lott 1987c, 1987d). Certainly Protestants benefited from the knowledge that public schooling was being used to instill young Catholics with beliefs these Protestants valued.

Since we observe governments all over the world going to great lengths to control the information their citizens receive, is it so surprising that governments would also try to control the information that their children receive through schooling? My empirical work shows that the variations in the public provision of television, radio stations, and education across countries operate in a very similar fashion with respect to such things as the level of transfers and the type of government, and suggests that controlling the media and education are close substitutes in producing the desired views in the population. Given that my theory hypothesizes that the return to indoctrination increases with both higher levels of transfers and totalitarianism (Lott 1990c), the recent changes in Eastern Europe over the last year should provide a dramatic test for my theory. If I

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

am correct, we should observe significant reductions in the level of public schooling expenditures in those countries.

References

Bender, Bruce. "A Reexamination of the Principal-Agent Relationships in Politics." Working Paper, University of Wisconsin at Milwaukee, 1990.

Bender, Bruce. "The Influence of Ideology on Congressional Voting." *Economic Inquiry* 29 (1991): forthcoming.

Coursey, Don L., and Roberts, Russell D. "Competition in Political and Economic Markets." *Public Choice* (1990): forthcoming.

Goff, Brian L., and Grier, Kevin B. "On the (Mis)Measurement of Legislator Ideology and Shirking." Working Paper, George Mason University, 1990.

Iannacone, Larry. "The Consequences of Religious Regulation." Hoover Institution Working Paper, presented at the 1990 Public Choice Association Meetings in Tucson, Ariz., 1990.

Lott, John R., Jr. "Political Cheating." *Public Choice* 52 (1987a): 169–86.

Lott, John R., Jr. "The Institutional Arrangement of Public Education: The Puzzle of Exclusive Territories." *Public Choice* 54 (1987b): 89–93.

Lott, John R., Jr. "Why Is Education Publicly Provided? A Critical Survey." *Cato Journal* 7 (Fall 1987c): 475–501.

Lott, John R., Jr. "Juvenile Delinquency and Education: A Comparison of Public and Private Provision." *International Review of Law and Economics* 7 (December 1987d): 163–75.

Lott, John R., Jr. "Attendance Rates, Political Shirking, and the Effect of Post-Elective Office Employment." *Economic Inquiry* 28 (January 1990a): 133–50.

Lott, John R., Jr. "Academic Freedom at a Public University." *The Freeman* 40 (March 1990b): 112–15.

Lott, John R., Jr. "An Explanation for the Public Provision of Schooling: The Importance of Indoctrination." *Journal of Law and Economics* 33 (April 1990c): 199–231.

Lott, John R., Jr., and Davis, Michael. "Ideological Shirking of Ideological Priors: A Critical Review and Extension of the Literature." Working Paper, Anderson Graduate School of Management, University of California at Los Angeles, 1990.

Peltzman, Sam. "The Growth of Government." *Journal of Law and Economics* 23 (October 1980): 209–87.

Peltzman, Sam. "Constituent Interest and Congressional Voting." *Journal of Law and Economics* 27 (1984): 181–210.

Van Beek, James. "Does the Decision to Retire Increase the Amount of Political Shirking?" Working Paper, Texas A&M University, 1990.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

BOOK REVIEWS

The Political Economy of the GATT
Peter Moser
Grusch, Switzerland: Verlag Ruegger, 1990, 141 pp.

According to Peter Moser, "The main purpose of international trade negotiations is not to trade off divergent interests between countries, but to modify the rules of interactions in the domestic political process." From this insight, Moser has made a major contribution to the growing literature on constitional political economy, among the first to be applied to international agreements.

Moser addresses the following questions, among others:

- What leads a national government to make trade agreements?
- What is the characteristic content of these agreements?
- How are such agreements enforced in the absence of a world government or some other common enforcement authority?
- What leads to the erosion of these agreements?
- What have been the effects of these agreements?
- What led Congress to authorize the president to make substantial reductions in tariffs in the Reciprocal Trade Agreements Act of 1934?

Part I of this monograph develops a general theory of the international economic order and then applies this theory to explain the characteristics of the General Agreement on Tariffs and Trade (GATT). Moser interprets trade agreements as serving the same objective as national constitutions—to restrain the domestic political process from actions that reduce the "general welfare" within each nation. Specifically, trade agreements are an exchange of commitments that attempt to constrain national policies that reduce access to foreign markets. In that sense, trade agreements become part of the effective economic constitution of each participating government. He argues convincingly that an open international economic order is not dependent on international enforcement but will emerge if national governments protect certain basic rights and if international contracts are enforced in the national courts. The observed pattern of trade restraints is the result of an *intra*national "prisoners' dilemma"

Cato Journal, Vol. 10, No. 1 (Spring/Summer 1990). Copyright © Cato Institute. All rights reserved.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

and not from divergent interests across countries. And, on the basis of the insight of the late Jan Tumlir, Moser concludes that the problems of the international economic order in this century are primarily a consequence of the erosion of the effective economic constitutions of the national governments.

Part II develops a model of the political equilibrium tariff rate and then tests this model with the history of U.S. tariffs from 1900 through 1986. The primary conclusion of this model is that unilateral tariff setting will lead to higher equilibrium tariffs than would result from a trade agreement. The model, however, does not lead to an unambiguous conclusion that national governments will seek a trade agreement. My only quibble with this model is that it assumes that a flow of mobile factors into the import-competing industries will increase the return in these industries; this increase would happen only if the mobile factors are complements, not substitutes, to the factors already employed in these industries. In the one case of an increase in a foreign tariff, the interests of labor and capital in the import-competing industries are likely to be different. It is not clear, however, that a different assumption in this case would change the primary conclusions of this model.

The empirical tests are generally supportive of the results of this model. From 1900 through 1939, average U.S. tariffs were strongly countercyclical, increased during a Republican presidency, and declined during World War I, but were not significantly affected by the Trade Act of 1934. During the GATT period through 1986, average tariffs were much lower and only weakly countercyclical but they increased strongly in response to the terms of trade; in addition, average tariffs were reduced substantially in response to the Trade Act of 1974 and the Kennedy and Tokyo Rounds, and were independent of the president's party. The more disturbing finding is that the number of affirmative decisions on antidumping petitions increased sharply in response to the decline in average tariffs.

Part III develops a model of congressional decisions on trade issues and then tests this model on the decision to approve the Reciprocal Trade Agreement Act of 1934. The model, based on the analysis by Barry Weingast, focuses on the critical role of the key congressional committees in providing durability to congressional decisions. The 1934 Act, however, does not provide a sufficient test of this model, because several political changes were consistent with the single fact of the approval of this Act: specifically, the election of a Democratic president, the election of a Democratic Senate, and the selection of members of the Senate Finance Committee who were more supportive of freer trade than was the median senator. A better test of this model would require an analysis of all of the major trade acts and trade agreements from 1934 through 1988.

On reading this monograph, I was most surprised to learn this is Moser's dissertation. It is a model dissertation and reflects much more

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

maturity than is characteristic of doctoral candidates. The combination of serious institutional analysis, enough mathematics to demonstrate his technical skills, careful empirical analysis, and clear writing (in Moser's second language) is most impressive. This monograph deserves a broad readership and should alert all of us to watch for Moser's future contributions.

William A. Niskanen

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

CRITICAL REVIEW

A JOURNAL OF BOOKS AND IDEAS

The interdisciplinary journal of classical liberal theory & research.

"Not content to simply repeat the ideas of the last two hundred years, or to ignore liberalism's critics, CRITICAL REVIEW has done a remarkable job of pushing liberalism forward by integrating the latest research in philosophy, economics and other social sciences and exposing it to critical reflection." — ***Market Process***, Spring 1989

"I am an avid reader of your excellent journal, the only new one that I have added to my list in recent years...."
— **Aaron Wildavsky**, University of California, Berkeley

In the current double issue:

Fukuyama: more right than he knew.

Why history "ended" long before Gorbachev, with the failure of central planning in 1921 and the convergence of East and West on the interventionist welfare state.

Also: challenges to the Keynesian orthodoxy

from John B. Egger, David Felix, Lowell Gallaway & Richard Vedder, Steven Horwitz, George Selgin, Frank Vorhies and Richard E. Wagner, disputed by Tyler Cowen and Paul Davidson.
Plus essays on socialism and Socrates.

Forthcoming: THE WELFARE STATE

debated by Harold Demsetz • Nathan Glazer • Antony de Jasay • Elisabeth Langby • Seymour Martin Lipset • Charles Murray • David L. Prychitko.

Plus: methodological individualism • communitarianism • environmentalism • jurisprudence • democracy • egalitarianism • the Third World...

Papers by: Norman Barry • David Beito • Stephan Boehm • Peter Boettke • Susan Love Brown • Gus diZerega • Antony Flew • Roger Garrison • John Gray • David Kelley • Israel Kirzner • Don Lavoie • Loren Lomasky • G.B. Madison • Martin Malia • Donald McCloskey • David Prychitko • Jeremy Shearmur • Barry Smith • Hernando de Soto • Andrzej Walicki...

Double issue $12. 1 year (4 issues): $24 individuals, $30 or £19 outside US; $39 libraries, $45 outside US. Add $15/yr. for foreign air mail. Send check, money order or Visa/MC number & exp. date to:
CRITICAL REVIEW, Dept. 201, P.O. Box 14528, Chicago, IL 60614, USA
ISSN 0891-3811

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

Reason Papers

A Journal of Interdisciplinary Normative Studies

Essays to Commemorate the 250th Anniversary of the Completion of David Hume's *A Treatise of Human Nature*

Edited by Stuart D. Warner

Articles

Hume on the Origin and Evolution of Religious and Philosophical Consciousness ..... *Donald W. Livingston* 3

The Virtue of Political Skepticism ..... *James T. King* 25

Hume's Account of Property ..... *Nicholas Capaldi* 49

David Hume on the Public Interest ..... *Stuart D. Warner* 77

Spinoza and Hume on Individuals ..... *Douglas Den Uyl and Lee Rice* 94

Natural Rights, Philosophical Realism, and Hume's Theory of Common Life ..... *Douglas B. Rasmussen* 122

Discussion Notes

In Defense of Moore's "Proof of an External World" ..... *John O. Nelson* 141

Race Isn't Merit ..... *Eugene Sapadin* 145

Book Reviews

William G. Scott and David K. Hart's *Organizational Values in America* ..... *David L Norton* 153

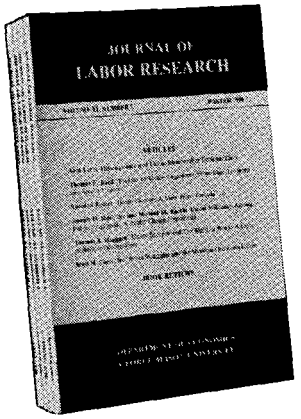
Hans-Hermann Hoppe's *A Theory of Socialism and Capitalism* ..... *Daniel Shapiro* 158

REASON PAPERS is published at the Department of Philosophy, Auburn University, AL 36849. Send orders ($7.00 per copy in US and Canada, $8.00 elsewhere; make checks payable to Reason Papers) to Professor Tibor R. Machan, *Reason Papers*, Department of Philosophy, Auburn University, AL 36849. Manuscripts should be accompanied with return postage and envelope.

Copyright©1990 by the Department of Philosophy, Auburn University. All rights reserved.

No. 15 **SUMMER 1990**

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED



JOURNAL OF LABOR RESEARCH

Selected Articles Forthcoming in Volume XI, 1990

SYMPOSIUM: Japanese Labor Relations: Are There Lessons for the U.S.?

The Community as a Source of Union Power	James A. Craft
Union Militancy Among Public Employees: A Public Choice Hypothesis	Joseph D. Reid, Jr. and Michael M. Kurth
Is the U.S. Unique in the Decline of Private Sector Unionism?	Leo Troy
The Changing Scope of Union Organizing	Gary N. Chaison and Dileep G. Dhavale
The Involvement of Canadian Unions in Political Activities: An Exploratory Analysis	John C. Anderson and John Thomas Delaney
Government Wage Differentials Revisited	William J. Moore and John Raisian
Monopoly, Efficient Contract, and Median Voter Models of Union Wage Determination: A Critical Comparison	Bruce E. Kaufman and Jorge Martinez-Vazquez
Union Decline: A View from Canada	Yonatan Reshef
Strike Strategies and the Minimum Effective Union	Scott M. Fuess, Jr.
Heterogeneity and Union Membership Determination	Anat Levy
Pensions and the Uses of Ignorance by Unions and Firms	Teresa Ghilarducci
Equity Market Response to Union Decertification Petitions and Elections	William L. Huth and Don N. MacDonald
Why People Work: An Examination of Interstate Variations in Labor Force Participation	Lowell Gallaway, Richard Vedder and Robert Lawson
Profiles of Union Organizers from Manufacturing and Service Unions	Thomas F. Reed
Discipline and Due Process in Labor Unions	J. Lawrence French and Jane K. Giacobbe
The Economic Effects of Faculty Unions	Randall G. Kesselring
The Effect of Voting Procedures on Union Representation Elections: The Multi-Union Case	Martha Norby Fraundorf

The **Journal of Labor Research** is published quarterly by the Department of Economics of George Mason University. Subscriptions are on a calendar year basis and subscribers will receive all issues for the current year. Annual subscription rates are: individuals, $30.00; academic libraries, institutions, government agencies, and business firms, $75.00. Foreign subscribers add $5.00 to cover additional postage. All remittances should be made payable to **Journal of Labor Research**, Department of Economics, George Mason University, Fairfax, Virginia 22030.

Editorial correspondence and manuscripts should be sent to James T. Bennett, Editor, Journal of Labor Research, Department of Economics, George Mason University, Fairfax, VA 22030. Manuscripts should be submitted in triplicate with a 100-word abstract. A submission fee of $20.00 for nonsubscribers must accompany all manuscripts. Papers accepted for publication must conform to JLR style requirements; style sheet available on request.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

JOURNAL OF MONEY, CREDIT, AND BANKING

Widely read and cited by researchers and policymakers, the *Journal of Money, Credit, and Banking* is a primary economics journal reporting major findings in the study of monetary and fiscal policy, credit markets, money and banking, portfolio management, and related subjects. Editors: Stephen A. McCafferty and J. Huston McCulloch, Ohio State University.

Recent and Forthcoming Articles:

Inflation and Taxation with Optimizing Governments
James M. Poterba and Julio J. Rotemberg

A Multicountry Characterization of the Nonstationarity of Aggregate Output
Robert C. Kormendi and Philip Meguire

Inflation and Reputation with Generic Policy Preferences *Herschel I. Grossman*

Loan Commitments and Optimal Monetary Policy
John V. Duca and David D. VanHoose

Policy Rules, Inflationary Bias, and Cyclical Stability
David M. Garman and Daniel J. Richards

Impact of Deposit-Rate Ceiling Changes on Bank Stock Returns *Haluk Unal*

Scottish Banking before 1845: A Model for Laissez-Faire?
Tyler Cowen and Randall Kroszner

The Risk Premium in the Foreign Exchange Market *Anne Sibert*

Published quarterly by the **Ohio State University Press**
1070 Carmack Road, Columbus, Ohio U.S.A. 43210

One-year subscription: $45.00 Library/Institution $27.50 Individual $20 Student

---

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

CONSTITUTIONAL POLITICAL ECONOMY

A New Interdisciplinary Journal
ISSN 1043-4062

Advisory Editor
JAMES M. BUCHANAN
Nobel Laureate

Editors
VIKTOR J. VANBERG
RICHARD E. WAGNER

"Yet, although the problem of an appropriate social order is today studied from the different angles of economics, jurisprudence, political science, sociology, and ethics, the problem is one which can be approached successfully only as a whole."

Friedrich A. von Hayek

The journal CONSTITUTIONAL POLITICAL ECONOMY is a forum for papers in the broad area of constitutional analysis, which lies at the intersection of several approaches in modern economics. All of these approaches share an interest in systematically integrating the institutional dimension—the study of political, legal, and moral institutions—into economic analysis.

Though its primary locus is in economics, CONSTITUTIONAL POLITICAL ECONOMY has an interdisciplinary orientation. One of its aims is to invite and encourage interdisciplinary exchange, including contributions from the various social sciences, philosophy, law, etc. Accordingly, CONSTITUTIONAL POLITICAL ECONOMY places particular emphasis on articles which, in style and substance, are suitable to this end. Theoretical papers and empirical studies are published, as well as contributions to constitutional policy issues. Reviews of pertinent books are also included.

To receive more information about CPE, write to the editorial office:

Constitutional Political Economy
Center for Study of Public Choice
George Mason University
Fairfax, Virginia 22030-4444

To subscribe to Vol. 1, Nos. 1-3, 1990, at the ***special introductory rate,*** enclose your check or money order with your subscription.

Please send information on CPE ☐
Please enter my individual subscription $20 ☐
Please enter my institutional subscription $50 ☐
(Subscribers outside the U.S. add $9 for air delivery or $6 for surface mail.)

Name _____ Institution _____
Street _____
City _____ State _____ Zip _____

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

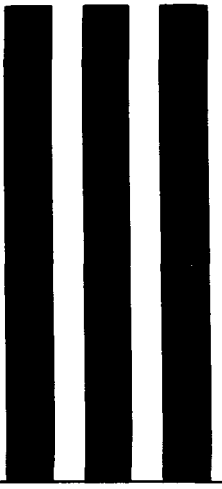
FUTURE MEETINGS
ATLANTIC ECONOMIC SOCIETY
CALL FOR PAPERS

WILLIAMSBURG, VIRGINIA
October 11-14, 1990

ROME, ITALY
March 15-22, 1991

PARTICIPANTS: Those preferring to be an author, chairman, or discussant must submit the following information: (1) Conference location; (2) Name; (3) Institution of affiliation; (4) Mailing address; (5) Phone number; (6) Preference for discussant, chairman, or both; (7) JEL number and primary and secondary interests. Authors must also enclose a $35.00 submission fee per paper. All accepted participants will be responsible for their own expenses and payment of the conference registration fee.

SUBMIT TO: John M. Virgo, Executive Vice President, Atlantic Economic Conference, Southern Illinois University at Edwardsville, Campus Box 1101, Edwardsville, IL 62026-1101 Phone: (618) 692-2291.



You can have it all...

What you're reading now plus the important articles in 800 other business and management magazines, in a matter of minutes.

The ABI/INFORM™ business database and your computer give you access to article summaries from magazines worldwide.

Call the publishers of ABI/INFORM at 800/626-2823, today.

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

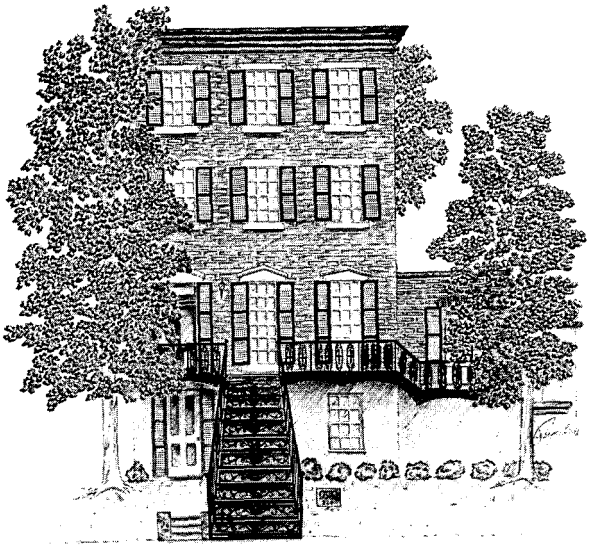
The Cato Institute

Founded in 1977, the Cato Institute is a nonpartisan public policy research foundation headquartered in Washington, D.C. The Institute is named for *Cato's Letters*, libertarian pamphlets that helped lay the philosophical foundation for the American Revolution.

The Cato Institute seeks to broaden the parameters of public policy debate to allow consideration of more options that are consistent with the traditional american principles of limited government, individual liberty, and peace. Toward that goal, the Institute strives to achieve a greater involvement of the intelligent, concerned lay public in questions of policy and the proper role of government.

The Cato Institute undertakes an extensive publications program dealing with the complete spectrum of policy issues. Books, monographs, and shorter studies are commissioned to examine the federal budget, Social Security, monetary policy, natural resource policy, military spending, regulation, NATO, international trade, and many other issues. Major policy conferences are held throughout the year, from which papers are published thrice yearly in the *Cato Journal*.

In order to maintain an independent posture, the Cato Institute accepts no government funding. Contributions are received from foundations, corporations, and individuals, and from the sale of publications. The Institute is a nonprofit, tax-exempt, educational foundation under Section 501(c)3 of the Internal Revenue Code.



The Watterston House headquarters of the Cato Institute

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED

$7.00

THE COMPOUND REPUBLIC

In a single republic, all the power surrendered by the people is submitted to the administration of a single government; and the usurpations are guarded against by a division of the government into distinct and separate departments. In the compound republic of America, the power surrendered by the people is first divided between two distinct governments, and then the portion allotted to each subdivided among distinct and separate departments. Hence a double security arises to the rights of the people. The different governments will control each other, at the same time that each will be controlled by itself.

—**James Madison**
Federalist No. 51

LICENSED TO UNZ.ORG
ELECTRONIC REPRODUCTION PROHIBITED